AIR POLLUTION ACCOUNTABILITY: ASSESSING REGULATORY IMPACTS ON EMISSIONS AND AIR QUALITY

A Dissertation
Presented to
The Academic Faculty

by

Lucas Richard Fath Henneman

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Civil and Environmental Engineering

Georgia Institute of Technology
August 2017

AIR POLLUTION ACCOUNTABILITY: ASSESSING AIR POLLUTION REGULATION IMPACTS ON EMISSIONS AND AIR QUALITY

Approved by:

Dr. Armistead G. Russell, Advisor
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Athanasios Nenes
School of Earth & Atmospheric Sciences
Georgia Institute of Technology

Dr. James A. Mullholland
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Paige E. Tolbert
Rollins School of Public Health
Emory University

Dr. Patricia L. Mokhtarian
School of Civil and Environmental Engineering
Georgia Institute of Technology

Date Approved: July 07, 2017

To Mom, Dad, and Nate.

ACKNOWLEDGEMENTS

This thesis could not have come together without the assistance of multiple people, to all of whom I am thankful. The most important: my advisor Dr. Armistead (Ted) Russell. His expertise, leadership, and connections were all invaluable to my work and to the amount I learned about air pollution accountability in my five years of graduate school.

The Accountability portion of this thesis was completed as part of a joint project with Emory University. Throughout my graduate school work, I benefited from discussions and feedback with Dr. Paige Tolbert, Dr. Mitch Klein, and Dr. Howard Chang. My counterpart at Emory, Dr. Joseph Abrams, gave me important insights into the logistics of epidemiological modeling.

I was supported by a network of excellent people at Georgia Tech. First, the Russell Group: Dr. Heather Holmes, Dr. Sunni Ivey, and Dr. David Lavoué all provided instrumental leadership during my first few years of graduate school. Dr. Cong Liu ran most of the meteorology, emissions, and air quality modeling I used for my research, and was a superb reference for indoor air quality and great office mate. I benefited from numerous discussions, group meetings, and social events with the rest of the group. Jessica Brown, of the environmental engineering program office, was an excellent resource for all things Georgia Tech. I very much enjoyed my time with the officers of the Georgia Tech Association of Environmental Engineers and Scientists (AEES).

Multiple stakeholders in the southeastern United States contributed their time and/or expertise in discussions, meetings, and emails. Dr. Gil Grodzinski and Jon Morton (of the Georgia Environmental Protection Division) both provided detailed information and data related to mobile source emissions modeling and the regulatory framework in Atlanta. My discussions with Steve Ewald and Charles Huling, formerly of Southern Company and Georgia Power, respectively, were instrumental in helping me wrap my head around how utilities respond to regulations and other drivers.

Funding is an important part of any research work. Mine came primarily from the National Science Foundation under the Graduate Research Fellowship Program Grant No. DGE-1148903 and the Health Effects Institute. This researched leveraged resources related to the Southeastern Center for Air Pollution and Epidemiology (SCAPE) under US EPA

grant R834799. This publications contents are solely the responsibility of the grantee and do not necessarily represent the official views of the US EPA. Further, US EPA does not endorse the purchase of any commercial products or services mentioned in the publication. The Georgia, Southern Section, and International Air and Waste Management Association chapters awarded me scholarships throughout my time as a PhD student, as did the Association of Environmental Engineers and Scientists at Georgia Tech.

The International Association for Applied Systems Analysis (IAASA) provided me an engaging opportunity to supplement my study at Georgia Tech with a three-month research stint in South Africa. I am grateful for all people who made that experience one that I'll never forget, particularly Dr. Carmen Klausbruckner, my research and travel companion, and my mentors, Drs. Peter Rafaj and Harold Annegarn.

I benefited from an active social network that periodically helped me escape the research life. My mental and physical states benefited immensely from experiences with friends, teammates, and the Chamblee High School Bulldog baseball team. I am grateful for two people in particular—Nick Stone, my friend of five years, and Laura Carpenter, my girlfriend. They shared in the good times and bore with grace the full brunt of disappointing results, late night brainstorms, and strict deadlines, and always had helpful ideas and positive thoughts to share in return.

Finally, I wouldn't have stood a chance at completing this work without my family. My mom (Dr. Janet Fath), dad (Dr. Richard Henneman), and brother (Nate) have given me every opportunity to succeed, and continue their generous support of my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

ANAA	Atlanta $PM_{2.5}$ non-attainment area
AQS	Air Quality System
CAAA	1990 Clean Air Act Amendments
CAIR	Clean Air Interstate Rule
CSAPR	Cross-State Air Pollution Rule
CTM	Chemical Transport Model
ED	Emergency Department
EGU	Electricity Generating Unit
EPA	United State Environmental Protection Agency
EPD	Georgia Environmental Protection Division
GAINS	Greenhouse Gas and Air Pollution Interactions and Synergies Model
GRAQC	Georgia Rules for Air Quality Control
IRP	Integrated Resource Plan
$KZ_{m,p}$	Kolmogorov-Zurbenko filter with window m and number of passes p
MDA8h	Maximum Daily 8-hour Average
MOB	Mobile-source
MOVES	EPA's Motor Vehicle Emissions Simulator
NAA	Non-Attainment Area
NAAQS	National Ambient Air Quality Standards
NBTP	NO_X Budget Trading Program
OPE	Ozone Production Efficiency
PGM	Photochemical Grid Model

$PM_{2.5}$	Particulate Matter with diameter less than 2.5 μm
PS*	emissions-independent atmospheric Photochemical State
PSC	Public Services Commission
REG	Regional, referring to EGU emissions in Alabama, Georgia, Mississippi, North Carolina, South Carolina, and Tennessee minus those in the ANAA
RH	Relative Humidity
SEARCH	Southeastern Aerosol Research and Characterization monitoring network
SIP	State Implementation Plan
VMT	Vehicle Miles Traveled

SUMMARY

The United States has seen large improvements in air quality over the last half century with the promulgation of regulatory actions designed to reduce air pollutant emissions to the atmosphere. Costs of these regulations, estimated by the United States Environmental Protection Agency (EPA) at tens of billions of dollars per year, motivate air pollution accountability research, which seeks to evaluate the effects of air quality regulations on emissions, air quality, exposure/dose, and public health—components of the so-called *Accountability Chain*.

Accountability research has grown as a field substantially since a 2003 Health Effects Institute Report defined the Chain and suggested best practices for Accountability studies. Since that time, studies have evolved from investigating short-term local pollution control actions to wide-ranging regulations implemented over multiple years. This thesis adds to the expanding scope in the field by investigating a range of regulatory actions on electricity generating units (EGUs, often called power plants) and on-road mobile sources promulgated in the United States since the 1990s. The 1990 Clean Air Act Amendments, passed by Congress to reduce emissions and subsequently improve air quality in the United States, serve as the basis for most of regulations reviewed here, though regulatory application at the state and local levels is complicated by the involvement of multiple regulating bodies. A detailed investigation herein shows that the United States has seen major decreases in emissions since the 1990s, and most of the decrease in emissions are attributable to regulatory policies, although other influences such as fuel costs, population demographic shifts, and technological improvements have influenced emissions reductions.

The bulk of the work described in this thesis investigates a central question for atmospheric researchers; namely, *what factors influence ambient air quality?* Specifically, the various chapters seek to link variability in meteorology and EGU/mobile source emissions to ambient air quality between the late 1990s and early 2010s. For example, detrending methods show that meteorology can have a large impact on daily air pollution concentrations, but removing this variability does not much impact annual or long-term trends for most pollutants in Atlanta, GA. Results of empirical statistical models and a

deterministic air quality model—both of which directly link EGU and mobile emissions estimates to observed air quality—show that emissions reduction programs have reduced the highest (summertime) ozone concentrations while simultaneously increasing the lowest (wintertime) concentrations. For particulate matter with diameter less than 2.5 μm ($PM_{2.5}$), controls have reduced both the annual mean values and the seasonality. These chapters establish important evidence that regulations were a primary driver of improved air quality, and that opportunities remain for further improving air quality through reduced emissions. Outcomes from these chapters was applied directly by colleagues at Emory University, who used the results to estimate that air quality regulations have led to 55,794 avoided cardiorespiratory emergency department visits in Atlanta.

Model evaluation, an important aspect of any model application, occupies a second major portion of this thesis. Results show that the deterministic model (the Community Multiscale Air Quality model (CMAQ), which is approved by the EPA for regulatory modeling) is adept at capturing ozone and $PM_{2.5}$ concentrations and changes over the decade. A deeper investigation, however, reveals that the model does not always get the right answer for the right reasons. For example, CMAQ simulates higher ozone production efficiency—OPE, a measure of the number of ozone molecules produced by anthropogenic emissions—than empirically derived OPE. The model has trouble estimating absolute concentrations and variability of certain species that make up $PM_{2.5}$—for example, the model is biased low for sulfate and organic species in the summer and biased high for organic species and elemental carbon in the winter—but the biases tend to cancel out when the species are summed.

The final body chapter maintains the focus on air quality policies, but projects their impacts into the future instead of assessing their effectiveness. Results from an integrated assessment model show that aggressive climate mitigation policies would have large air quality benefits, but air pollution policies are still necessary to reduce emissions important to human health, such as for heating and cooking in the domestic sector.

This research provides important evidence that causally links regulations to emissions reductions and air quality improvements while accounting for numerous concurrent changes. The accountability research undertaken hear represents the most detailed approach to date at tracking the impacts of air quality regulations implemented

over a long period across a broad spatial scale. The approach and results, which are detailed in a report soon to be published by the Health Effects Institute and multiple refereed journal articles, provide important evidence that air pollution regulations have indeed been effective at reducing emissions, improving air quality, and reducing incidences of negative public health outcomes. The final chapter discusses conclusions and future research, split into three sections: accountability, air quality modeling, and policy projections.

CHAPTER 1. INTRODUCTION

The United States has seen large improvements in air quality over the last half century with the promulgation of regulatory actions designed to reduce air pollutant emissions to the atmosphere (National Research Council 2004). These regulations, though viewed as successful at reducing anthropogenic contributions to local, regional, and global pollution, have come with major costs, estimated at $53 billion per year in 2010 (2006$) and expected to increase to $65 billion per year in 2020 (U.S. EPA 2011). These costs motivate air pollution accountability research, which seeks to evaluate the effectiveness of air quality regulations on emissions, air quality, exposure/dose, and public health—components of the so-called *Accountability Chain* (Health Effects Institute 2003). Public health, the primary motivator of efforts to improve air quality, similarly drives accountability research. Conclusions on the effectiveness of regulatory programs at improving public health, however, typically are built upon evidence of how regulations impacted the previous links in the Chain. This thesis seeks to establish that evidence and methods to develop such evidence more generally.

This thesis investigates a range of regulatory actions on electric generating units (EGUs, often called power plants) and on-road mobile sources promulgated in the United States since the 1990s. The 1990 Clean Air Act Amendments, passed by Congress to reduce emissions and subsequently improve air quality in the United States, serve as the basis for most of regulations reviewed here, though regulatory application at the state and local levels is complicated by the involvement of multiple regulating bodies (National Research Council 2004).

This work describes multiple approaches for propagating impacts of regulations, separate the impacts from other (i.e., non-regulatory) influences, and assess uncertainties at each relevant step in the Accountability Chain (Figure 1-1). Many of the chapters, therefore, build off one another by incorporating results from chapters at previous links in the Chain.

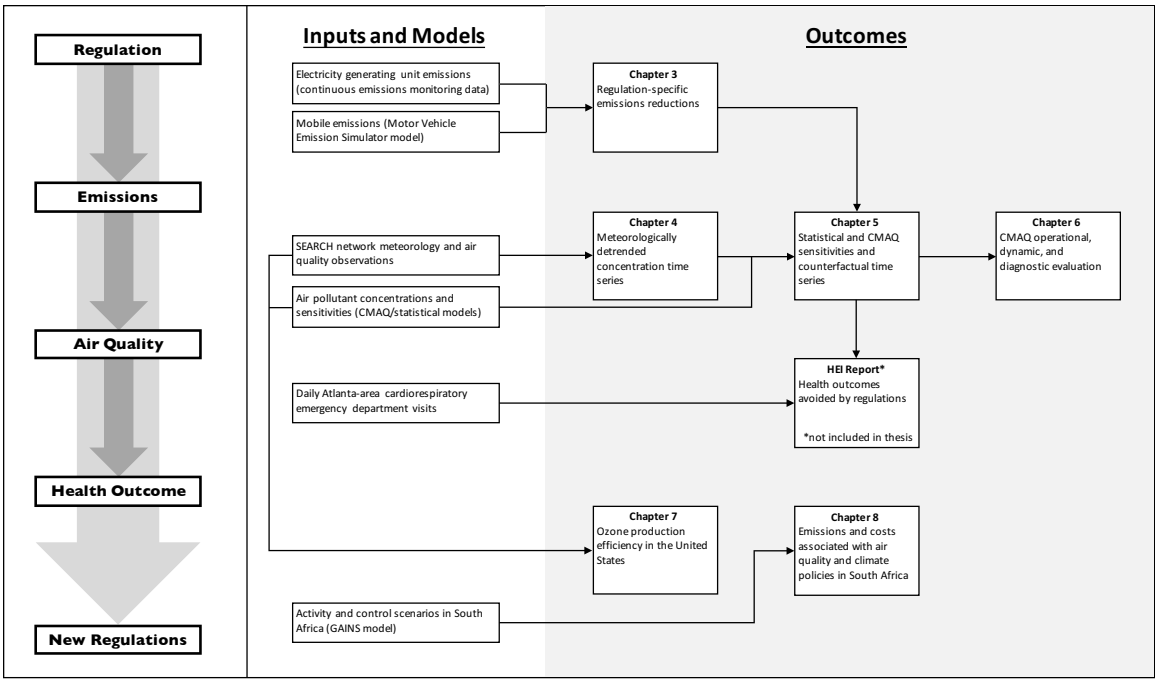


Figure 1-1. Primary paths of information flow through the work in this thesis, and each chapter's positions along the (slightly modified) Accountability Chain (left). The Health Effects Institute (HEI) report details the health outcomes research led by colleagues at Emory University, and is not reproduced here. Acronyms: SEARCH, the Southeastern Aerosol Research and Characterization monitoring network; CMAQ, the Community Multiscale Air Quality model; GAINS, the Greenhouse Gas and Air Pollution Interactions and Synergies Model.

The second chapter provides an in-depth the history of air quality accountability research. Early accountability studies focused on interventions implemented at specific dates (such as the ban of the use of coal for heating in Dublin, Ireland (Clancy et al. 2002) or restrictions on vehicular travel instituted during the Olympic games in Atlanta, GA and Beijing, China (Friedman et al. 2001; Rich et al. 2015)). Later studies expanded their scope to include regulations implemented over wider spatial and temporal scales, and researchers have employed new methods to causally attribute changes along the accountability chain to regulations while accounting for concurrent changes in other factors. The chapter concludes with recommendations for accountability study design.

The third chapter presents a detailed assessment of major EGU and mobile source regulations promulgated since the 1990 Clean Air Amendments, and benefitted from discussions with regulators and other stakeholders in the southeastern United States. Between the mid-1990s and 2013, EGU NO_X ($NO + NO_2$) and SO_2 emissions each decreased by about 80%, while mobile source emissions are estimated to have decreased

between 61% and 93% across a range of pollutants. Most of the decrease in emissions are attributable to policies, although other influences, such as fuel costs, population demographic shifts, and technological improvements influence emissions reductions. Unique methods that investigate changes in daily emissions factors allow for the resolution of impacts of regulations implemented incrementally over multiple years. Complicated relationships between federal and state regulatory policies and impacts of fuel prices (e.g., the emergence of cheap natural gas) on regulatory effectiveness, however, are more difficult to quantify.

The fourth chapter is the first in a series of chapters that characterize influencing factors of ambient air quality. In this chapter, multiple statistical detrending methods are applied to estimate daily impacts of meteorological variability on observed concentrations in Atlanta, GA—a process called *meteorological detrending*. Previous efforts at applying meteorological detrending have focused on ozone (O_3), a pollutant that is regulated to protect public health and is formed in the atmosphere. This work extended the detrending to multiple gaseous and particulate pollutants that have been linked to public health and negative environmental outcomes. Results show that on meteorology can have a large impact on daily concentrations, but removing this variability does not much impact annual or long-term trend for most pollutants.

The fifth chapter extends this work with the application of statistical models that directly link EGU and mobile emissions estimates to observed air quality in Atlanta. These models were evaluated beyond the typical statistical fit metrics by comparing pollutant sensitivities to emissions with similar values estimated using the Community Multiscale Air Quality (CMAQ) Model, a chemical transport model. Combining the statistical models with regulatory-impacted emissions changes determined in Chapter 3 resulted in counterfactual—i.e., assuming no further regulations after the 1990 Clean Air Act Amendments—O_3 and $PM_{2.5}$ (particulate matter with diameter less than 2.5 μm) concentrations. Results show that emissions reduction programs have reduced the highest (summertime) O_3 concentrations while simultaneously increasing the lowest (wintertime) concentrations. For $PM_{2.5}$, controls have reduced both the annual mean values and the seasonality.

Results of Chapters 3-5 were used directly by colleagues at Emory University to estimate the numbers of Atlanta-area cardiorespiratory emergency department (ED) visits avoided by air quality regulations (Russell et al., 2017). Using counterfactual pollutant time series and their uncertainties, they concluded that 55,794 cardiorespiratory ED visits were avoided due to air quality regulations. Annual avoided ED visits increased over time as regulations were phased in to their full implementations; the later years yielded the greatest benefit of the regulations.

CMAQ, an important regulatory decision-making tool, was applied in the sixth chapter to estimate emissions-air quality relationships and assess their changes over time. Operational, dynamic, and diagnostic evaluations were performed to assess the model's ability to capture changes in ambient concentrations and concentration sensitivities to emissions and meteorology. The work identifies important areas in which CMAQ is proficient and deficient at getting the correct answers for the correct reasons. CMAQ does well estimating the change in annual total $PM_{2.5}$ between 2001 and 2011, but has more difficulty estimating changes in certain species, such as sulfate and elemental carbon. A more detailed comparison with the statistical models applied in Chapter 5 reveal that the two models are similarly sensitive to emissions and meteorology changes across the decade for O_3 and $PM_{2.5}$, but important differences exist in $PM_{2.5}$ species sensitivities. This portion of the work informs both future improvements in CMAQ parameters and inputs, and policy decisions based on model outputs.

The seventh chapter delves deeper into how atmospheric chemistry has been impacted by reduced emissions. Analysis focuses on ozone production efficiency (OPE—a measure of the number of ozone molecules produced by each emitted NO_X ($NO_X = NO + NO_2$) molecule. Regulatory models such as CMAQ must correctly characterize OPE to support decisions on future approaches to reduce O_3 concentrations. Since OPE increases as available reactive nitrogen in the atmosphere decreases, it is important to evaluate OPE at low NO_X levels to ensure benefits of emissions reductions are not outweighed by increasing NO_X efficiency at producing O_3 in the atmosphere. Empirical finds that, while OPE has increased with emissions reductions, there is a spatially varying upper limit. CMAQ captures these changes relatively well, though slightly underestimates model responsiveness to emission changes. Overall, this evidence suggests

that NO_X emissions reductions remain a viable solution for reducing the highest O_3 concentrations.

The seventh chapter presents an evaluation of air quality policies, with a focus on the projection of regulatory effects on emissions into the future and the benefits and tradeoffs associated with various air quality and climate mitigation policies. Results from the Greenhouse Gas and Air Pollution Synergies model (GAINS) show that aggressive climate mitigation policies would have large air quality benefits, but air pollution policies are still necessary to reduce emissions important to human health, such as for heating and cooking in the domestic sector.

The final chapter discusses conclusions and future research, split into three sections: accountability, air quality modeling, and policy projections. Future accountability studies should leverage advanced planning and multidisciplinary teams to develop and evaluate methods that link regulations to various outcomes at each link in the Accountability Chain. Researchers should employ expanded observational abilities (e.g., low-cost sensors, expansive monitoring networks, and satellites) to evaluate and improve air quality models. Further, air quality model evaluation and application will benefit from new data handling and statistical techniques. Policy projections will occur in different regimes for developed and developing countries. Developed countries will be increasingly concerned with background air pollution levels as their own contributions decline, and developing countries will be able to leverage the experience of developed countries. Projections in both groups will need to account for added uncertainty associated with a changing climate.

This research, overall, shows what is already established: that regulations have improved air quality. However, the in-depth analysis provides important evidence that causally links regulations to emissions reductions and air quality improvements while accounting for numerous concurrent changes. It shows that the efficacy of past regulations has been variable: some provided substantial benefits while others have had limited effects. The work informed work by colleagues at Emory University who estimated the negative health effects avoided by the regulations described in detail here, thus completing the accountability study along the length of the Chain. This metered approach to following regulatory impacts across all of the links in the Accountability Chain will serve as an

example for future policy assessments. The bulk of the research is published in refereed journal articles, and is summarized in detail in a report to the Health Effects Institute (Russell et al., 2017).

1.1 References

Clancy, Luke, Pat Goodman, Hamish Sinclair, and Douglas W Dockery. 2002. "Effect of Air-Pollution Control on Death Rates in Dublin, Ireland: An Intervention Study." *Lancet* 360 (9341): 1210–14. doi:10.1016/S0140-6736(02)11281-5.

Cropper, Maureen, Arthur Fraas, and Richard Morgenstern. 2017. "Looking Backward to Move Forward." *Science* 355 (6332): 1375–76. http://www.sciencemagazinedigital.org/sciencemagazine/31_march_2017?sub_id=CDdH64IJVjiS&u1=01469606&folio=1375&pg=31#pg31.

de Hoogh, Kees, Michal Korek, Danielle Vienneau, Menno Keuken, Jaakko Kukkonen, Mark J. Nieuwenhuijsen, Chiara Badaloni, et al. 2014. "Comparing Land Use Regression and Dispersion Modelling to Assess Residential Exposure to Ambient Air Pollution for Epidemiological Studies." *Environment International* 73. Elsevier Ltd: 382–92. doi:10.1016/j.envint.2014.08.011.

Fiore, Arlene M., Vaishali Naik, Eric M. Leibensperger, and John D. Bachmann. 2015. "Air Quality and Climate Connections." *Journal of the Air & Waste Management Association* 65 (6): 645–85. doi:10.1080/10962247.2015.1040526.

Friberg, Mariel D., Xinxin Zhai, Heather A. Holmes, Howard H. Chang, Matthew J. Strickland, Stefanie Ebelt Sarnat, Paige E. Tolbert, Armistead G. Russell, and James A. Mulholland. 2016. "Method for Fusing Observational Data and Chemical Transport Model Simulations to Estimate Spatiotemporally Resolved Ambient Air Pollution." *Environmental Science and Technology* 50 (7): 3695–3705. doi:10.1021/acs.est.5b05134.

Friedman, MS, KE Powell, L Hutwagner, L M Graham, and W G Teague. 2001. "Impact of Changes in Transportation and Commuting Behaviors during the 1996 Summer Olympic Games in Atlanta on Air Quality and Childhood Asthma." *Jama* 285 (7): 897–905. http://www.ncbi.nlm.nih.gov/pubmed/11180733.

Health Effects Institute. 2003. "Assessing Health Impact of Air Quality Regulations: Concepts and Methods for Accountability Research." http://pubs.healtheffects.org/getfile.php?u=261.

Henneman, Lucas RF, Howard H Chang, David Lavoue, James A Mulholland, and Armistead G Russell. 2017. "Accountability Assessment of Regulatory Impacts on Ozone and PM2.5 Concentrations Using Statistical and Deterministic Pollutant Sensitivities." *Air Quality Atmosphere and Health*. doi:10.1007/s11869-017-0463-2.

Henneman, Lucas RF, Heather A Holmes, James A Mulholland, and Armistead G Russell. 2015. "Meteorological Detrending of Primary and Secondary Pollutant Concentrations: Method Application and Evaluation Using Long-Term (2000-2012) Data in Atlanta." *Atmospheric Environment* 119. Elsevier Ltd: 201–10. doi:10.1016/j.atmosenv.2015.08.007.

Henneman, Lucas RF, Cong Liu, Yongtao Hu, James A Mulholland, and Armistead G Russell. n.d. "Air Quality Modeling for Accountability Research: Operational, Dynamic, and Diagnostic Evaluation."

Henneman, Lucas RF, Cong Liu, James A Mulholland, and Armistead G Russell. 2016. "Evaluating the Effectiveness of Air Quality Regulations: A Review of Accountability Studies and Frameworks." *Journal of the Air & Waste Management Association* 67 (2). Taylor & Francis: 144–72. doi:10.1080/10962247.2016.1242518.

Henneman, Lucas RF, Peter Rafaj, Harold J. Annegarn, and Carmen Klausbruckner. 2015. "Assesing Emission Levels and Costs Associated with Climate and Air Pollution Policies in South Africa." *Energy Policy* 89. Elsevier: 160–70. doi:10.1016/j.enpol.2015.11.026.

Hu, Y., S. Balachandran, J. E. Pachon, J. Baek, C. Ivey, H. Holmes, M. T. Odman, J. A. Mulholland, and A. G. Russell. 2014. "Fine Particulate Matter Source Apportionment Using a Hybrid Chemical Transport and Receptor Model Approach." *Atmospheric Chemistry and Physics* 14 (11): 5415–31. doi:10.5194/acp-14-5415-2014.

Kloog, Itai, Francesco Nordio, Brent a Coull, and Joel Schwartz. 2012. "Incorporating Local Land Use Regression And Satellite Aerosol Optical Depth In A Hybrid Model Of Spatiotemporal PM2.5 Exposures In The Mid-Atlantic States." *Environmental Science & Technology*. doi:10.1002/dev.21214.Developmental.

Kwok, R. H F, K. R. Baker, S. L. Napelenok, and G. S. Tonnesen. 2015. "Photochemical Grid Model Implementation and Application of VOC, NOx, and O3 Source Apportionment." *Geoscientific Model Development* 8 (1): 99–114. doi:10.5194/gmd-8-99-2015.

National Academies of Sciences Engineering and Medicine. 2016. *The Future of Atmospheric Chemistry Research: Remembering Yesterday, Understanding Today, Anticipating Tomorrow*. Washington, DC: The National Academies Press. doi:10.17226/23573.

National Research Council. 2004. *Air Quality Management in the United States*. National Academies Press. doi:10.17226/10728.

Oreskes, Naomi, Kristin Shrader-Frechette, and Kenneth Belitz. 1994. "Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences." *Science* 263 (5147): 641–46.

Rich, David Q., Kaibo Liu, Jinliang Zhang, Sally W. Thurston, Timothy P. Stevens, Ying Pan, Cathleen Kane, et al. 2015. "Differences in Birth Weight Associated with the 2008 Beijing Olympic Air Pollution Reduction: Results from a Natural Experiment." *Environmental Health Perspectives* 117 (11): 1713–17. doi:10.1289/ehp.1408795.

Russell, Armistead G, Paige E Tolbert, Lucas RF Henneman, Joe Abrams, Cong Liu, Mitch Klein, James A Mulholland, et al. 2017. "Impacts of Regulations on Air Quality and Emergency Department visits in the Atlanta metropolitan area, 1999-2013" *In press*.

U.S. EPA. 2011. "The Benefits and Costs of the Clean Air Act from 1990 to 2020 Final Report."

CHAPTER 2. EVALUATING THE EFFECTIVENESS OF AIR QUALITY REGULATIONS: A REVIEW OF ACCOUNTABILITY STUDIES AND FRAMEWORKS[1]

2.1 Abstract

Assessments of past environmental policies—termed accountability studies—contribute important information to the decision-making process used to review the efficacy of past policies, and subsequently aid in the development of effective new policies. The Health Effects Institute defines the air pollution accountability framework as a chain of events that includes the regulation of interest, air quality, exposure/dose, and the health outcome of interest, and suggests that accountability research should address impacts for each of these linkages. Early accountability studies investigated short-term, local regulatory actions (for example, coal use banned city-wide on a specific date or traffic pattern changes made for Olympic games). More recent studies have switched to regulations made over longer time and larger spatial scales. Studies on broader scales require accountability research methods that account for effects of confounding factors that vary over time and space. Improved estimates of appropriate baseline levels (sometimes termed 'counterfactual'—the expected state in a scenario without an intervention) that account for confounders and uncertainties at each link in the accountability chain will help estimate causality with greater certainty. In the *direct* accountability framework, researchers link outcomes with regulations using statistical methods that bypass the link-by-link approach of classical accountability. Direct accountability results and methods complement the classical approach. New studies should take advantage of advance planning for accountability studies, new data sources (such as satellite measurements), and new statistical methods. Evaluation of new methods and data sources is necessary to improve investigations of long-term regulations, and associated uncertainty should be accounted for at each link to provide a confidence estimate of air quality regulation

[1] A modified version of this chapter has been published as: Henneman, L.R.F., Liu, C., Mulholland, J., Russell, A. (2017) "Evaluating the effectiveness of air quality regulations: A review of accountability studies and frameworks." *Journal of the Air and Waste Management Association*. Vol. 67, Issue 2

effectiveness. The final step in any accountability is the comparison of results with the proposed benefits of an air quality policy.

Implications: The field of air pollution accountability continues to be important to a number of stakeholders. Two frameworks, the classical accountability chain and direct accountability, have been used to estimate impacts of regulatory actions, and both require careful attention to confounders and uncertainties. Researchers should continue to be develop and evaluate both methods as they investigate current and future air pollution regulations.

2.2 Introduction

Air pollution accountability is a growing field that seeks to assess the effectiveness of regulatory actions, with governments and stakeholders that must control their emissions taking particular interest in the outcomes of such studies (Health Effects Institute, 2003). Early work in the field focused on air pollution and health impacts of local regulatory or non-regulatory actions (such as a major local air pollution source shutting down, e.g., due to economic factors). In the last decade, the methods have expanded to include long-term assessments of multiple control programs implemented over large spatial domains.

2.2.1 The Accountability Chain

The later decades of the 20th century saw major reductions in ambient air pollution concentrations in developed nations as governments enacted regulatory actions to curb emissions and subsequently protect human and environmental health (Hubbell et al., 2009; National Research Council, 2004). These three elements–regulations, emissions, and health–represent three of the five links in the *Accountability Chain* (air quality and exposure are the others: Figure 2-1, left side), which tracks the effects of an air quality regulation from promulgation to final impacts on public health. The goal of accountability research is to assess environmental policies by attributing causality (or lack thereof) of changes to links in the accountability chain to a regulatory action, and a successful accountability study should address each link in the chain between the regulatory action and the endpoint of interest (Health Effects Institute, 2003). At the conclusion of the analysis, study results should be compared with estimates of the proposed benefits of the regulations to air policy makers in decisions regarding future regulations. The majority of studies completed to present define either air pollution concentrations or health outcomes

as the endpoint of interest, while a few also assess economic impacts. Numerous studies have been published in this domain since the HEI's 2003 report (e.g. Dockery et al., 2013; Kelly et al., 2011; Morgenstern et al., 2012; Peel et al., 2010; Peters et al., 2009; Wong et al., 2012).

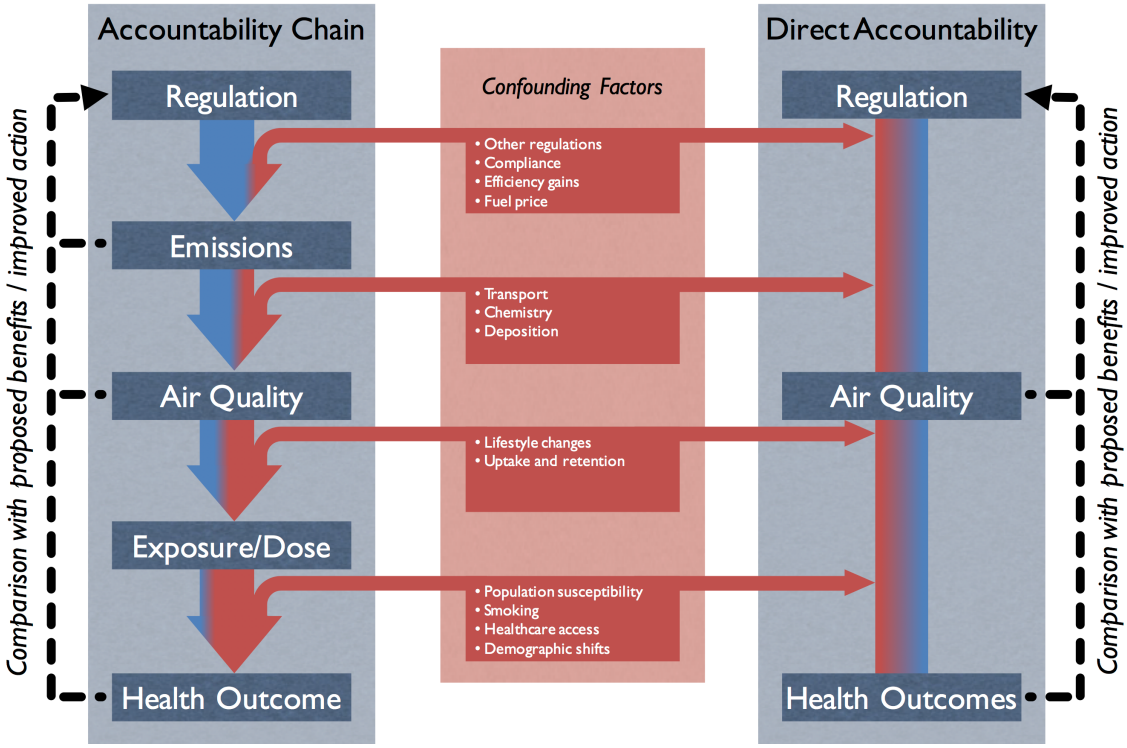


Figure 2-1. Classic accountability chain (left) and direct accountability framework (right). Confounding factors (red portion of the arrows) increasingly obscure the observable signals attributable to regulations (blue portion) between the links in the accountability chain. Further down the chain, the observed response is impacted only partially by regulations. Direct accountability studies seek to directly link regulatory actions to changes in air quality and/or health outcomes. Classic and direct accountability studies strive to take into account as many confounders as practical, though they use different methods to do so. At multiple points in both frameworks, the opportunity exists to compare with each regulation's proposed benefits and update the approach in new regulatory actions.

Defining causal relationships between any one of the links presents a difficult challenge, and assessing each link in the chain is even more arduous (Bell et al., 2011). Often, studies link changes in the outcome of interest to regulatory actions by making assumptions about the intermediate links in the accountability chain. Most studies identify a regulatory action of interest, and then use statistical or deterministic modeling of multiple other links in the chain to associate emissions with air quality and/or air quality with a

health outcome (e.g., Clancy et al., 2002; Hou et al., 2010; Kelly et al., 2011; Morgenstern et al., 2012; U.S. EPA, 2011, 1999a; Zigler et al., 2012).

The Health Effects Institute (2003) noted that a successful accountability study should address effects of regulatory actions on outcomes at each link in the accountability chain. Confounders at each link cloud the signal, and these factors compound through the length of the chain, increasingly obscuring the signal from the regulation at each link (Figure 2-1). Further, confounding variables and changes due to a regulatory action may not be wholly independent. For example, electric utilities use a complex decision-making process when deciding whether to retrofit or shutter old plants and/or build new ones. Inputs to these decisions include regulations, projected population changes and fuel costs, and geographic constraints, and can include stakeholder review and public oversight (National Research Council, 2004). This multitude of variables that affects these decisions makes it impossible to attribute some changes in emissions wholly to specific regulations. Given the issues involved in approaching accountability link-by-link, Zigler and Dominici (2014) suggest that direct accountability (discussed in more detail below) provides an attractive alternative.

Regulations. The first step in the accountability chain is the adoption of a policy and the ensuing regulations or related interventions. The regulation may or may not call for specific changes in emissions, and is, itself, subject to varying levels of success; e.g., the degree to which the regulation leads to actions that could result in emissions changes. For example, an area being found to be in non-attainment of the National Ambient Air Quality Standards in the United States does not lead directly to a set of specific emissions controls—each state that contains a non-attainment area submits a State Implementation Plan that details the actions it will take to reduce pollution levels (National Research Council, 2004; Zigler et al., 2016). These actions could include a regulation that specifies legal emissions amount and/or controls on a facility. An example of this in Georgia State—the Multipollutant Control Rule promulgated in 2008lists specific control technologies for each unit of electric generating units in the Atlanta area, which was in non-attainment for both ozone and $PM_{2.5}$ (Georgia EPD, 2013). Regulations may also lead to or be associated with unintended consequences, such as increased vehicle traffic carrying visitors to the Shanghai World Expo during a period of reduced industrial activity (Hao et al., 2011).

Emissions. Emissions of air pollutants are affected by multiple factors at a range of time scales from decadal to less than daily, including climate, population growth/decline, changes in the economy (including economic effects of air quality emissions, costs to operate in different areas, and fuel prices), regulatory actions, meteorology, weekly demand, and work travel patterns. Because of these factors and the general difficulty and expense of measuring emissions, typically only large sources, such as power plant stacks, are directly measured. Emissions from other sources (e.g. automobiles) are estimated as the sum of multiple smaller or individual sources. It is not yet practical to measure emissions from all of the cars all of the time, so the total fleet emissions are modeled, which introduces uncertainties (U.S. EPA, 2011). Uncertainties in data used by those models include emissions factors and activities, and studies have identified major biases in emissions estimates. For example, in the 1980s, the SCAQS tunnel studies identified that VOC and CO emissions were much higher than estimated (Lawson, 1990; Pierson et al., 1990), and that a small number of vehicles emitted a high percentage of total mobile emissions (these vehicles were termed 'super-emitters') (Lawson et al., 1990). Current ambient ground- and satellite-based observations studies have found a disagreement of up to a factor of two in emissions modeling (Anderson et al., 2014; Goldberg et al., 2016; Souri et al., 2016).

The lack of direct measurements makes it challenging to assess how emissions are changing, much less how they are responding to specific regulations. Further, in developed and developing countries alike, emissions sources are regulated under a number of statutes implemented at various times and applied over various spatial scales (Health Effects Institute, 2003). Such complexities are likely to increase as interest grows in a multimedia approach to environmental issues (Klausbruckner et al., 2016; Murray, 2013).

Even for sources whose emissions are well characterized, it can be difficult to assess the impacts of regulations. For example, if a regulation would lead to costly controls or plant retrofits, a utility may decide to shutter a plant before they otherwise would, "move" emissions, such as by increasing activities in one location and decreasing activities in another, or, in order to avoid New Source Review requirements, keep older plants online (National Research Council, 2004). Other external factors can play a role, such as economic activity that drives reduction in demand during recessions (Chay and

Greenstone, 2003) or changing fuel costs, which may or may not be directly linked to regulatory actions (National Research Council, 2004). Quantifying the actual impacts becomes more uncertain as the length of time assessed grows.

Studies have employed various techniques for addressing each of these challenges in order to attribute changes in emissions to specific interventions. For example, estimates of how mobile emissions have changed in response to regulations come from studies on a limited number of vehicles, changes in emissions factors and ambient observations (e.g., in tunnels), and emissions models (McDonald et al., 2013; National Research Council, 2004). In studies on interventions that occurred over short time periods, for instance, long-term trends in emissions are typically ignored, and baseline emissions are taken as the period before (or after, in the case of interventions that expire) the intervention (Friedman et al., 2001; Hou et al., 2010; Peel et al., 2010). Studies on long-term interventions typically either assess changing emissions over time (e.g., Butler et al., 2011; Gégo et al., 2007; Godowitch et al., 2010) or model an estimate of counterfactual—i.e. what would have happened had there not been an intervention—by assuming a constant emissions rate from a baseline year throughout the study period (e.g., Daskalakis et al., 2016; Gégo et al., 2008; Morgenstern et al., 2012).

Air Quality. The impacts of regulations and emissions on air quality have been well studied using both empirical (Blanchard et al., 2012; Butler et al., 2011; Harrington et al., 2012; Pierce et al., 2010; Zigler et al., 2012) and non-empirical (i.e. deterministic) air quality modeling (e.g., Daskalakis et al., 2016; Simon et al., 2015; Wang et al., 2014) approaches, each of which has strengths and limitations. Empirical approaches directly use observations of air quality, thus are tied to observations that are presumed to have relatively small uncertainties. Statistical techniques provide both associations between emissions changes and air quality and uncertainty estimates. However, statistical linkages to emissions changes or the underlying interventions are subject to confounding factors and require a number of modeling assumptions. If not all factors are correctly accounted for and assumptions justified in the approach, the associations will be flawed (Chang et al., 2014; Harrington et al., 2012; Zigler et al., 2012). Deterministic air quality models include state of the science characterizations of atmospheric processes, but are subject to

uncertainties in model inputs and parameterizations and numerical errors (Napelenok et al., 2011; Simon et al., 2012; Appel et al., 2008).

Two empirically-based approaches commonly used for evaluating the impacts of regulations on air quality are time series analysis and source apportionment. Time series analysis involves relating changes in meteorology and/or emissions to ambient measurements. Methods that link only meteorological variability (not emissions) to ambient concentrations are focused on detrending air pollution in order to assess long-term trends without added interference of meteorology (Camalier et al., 2007; Cox and Chu, 1993; Flaum et al., 1996; Gégo et al., 2007; Henneman et al., 2015; Kuebler et al., 2001; Milanchus et al., 1998; Rao et al., 1995). Other methods seek to directly relate changes in measured concentrations with regulatory actions (Zigler et al., 2012), emissions changes (Butler et al., 2011; Harrington et al., 2012), or fuel consumption (Russell, 2015). Such methods can be used with counterfactual emissions estimates to model potential outcomes that assume no regulations. Source apportionment (the attribution of observed concentrations to specific sources) remains an active area of research (Blanchard et al., 2012; Hopke, 2016; Hu et al., 2014; Yuan et al., 2013). Empirical accountability studies typically define a baseline period of air quality during the time periods before and after an intervention (e.g., Clancy et al., 2002; Friedman et al., 2001; Peters et al., 1996), or, alternatively, in the location of the intervention versus a control region (e.g., Lin et al., 2012; Peel et al., 2010; Wong et al., 1998), and quantify resulting changes. Zigler and Dominici (2014) point out potential issues with the classical approach, including the fact that it relies on extrapolating associative relationships and counterfactual scenarios that cannot be evaluated using measurements.

Deterministic air quality models directly estimate the impacts of emissions changes on changes in air quality. Such models make it possible to distinguish the effects of multiple regulations happening simultaneously through the estimation of emissions-concentration sensitivities (Napelenok et al., 2011; Simon et al., 2013). These models have been used to estimate counterfactual concentrations (Daskalakis et al., 2016; Gégo et al., 2008; Godowitch et al., 2010; Wang et al., 2014) and are primary tools used to assess proposed regulations (e.g., U.S. EPA, 2005a, 2005b, 1997a) and demonstrate future

attainment with air quality standards as part of State Implementation Plans (SIPs) (National Research Council, 2004).

Exposure/Dose. Challenges that arise in estimating exposure (and the subsequent health effects) to air pollution in accountability research are similar to those in other air pollution epidemiologic research (Cox and Popken, 2015). In most cases, scientists use air quality at a central monitor (or a group of monitors) to approximate exposure for all cases in the sample population, though recent studies have used results from combined multiple models and satellite data (e.g., Lee et al., 2015; Pachon and Balachandran, 2012). For daily statistical time series models, researchers often include multiple lag terms or averaging periods to assess various relationships between health outcomes and measured concentrations. A few studies use population-weighted concentration (Hou et al., 2010) and personal exposure monitoring (Wu et al., 2010). Some studies opt to not assess exposure, and instead apply statistical techniques that test for a change in health outcomes rates over the pre-defined time period of the intervention (Greenstone, 2004; Pegues et al., 2012; Zigler et al., 2016). Specific issues in relating air quality changes to exposures when conducting accountability studies include diagnosing both model uncertainty and parameter uncertainty in model selection (Morgenstern et al., 2012), missing data (Bell et al., 2011; van Erp et al., 2011), and estimating baseline exposure for the no-control case that accounts for appropriate confounders (Health Effects Institute, 2003).

Health Effects. Increased air pollution has been linked to negative health outcomes in numerous studies (Dockery et al., 1993; Garcia et al., 2011; Laden et al., 2006; Lim et al., 2013; S. Lin et al., 2013). Outcomes of greatest interest include morbidity and mortality associated with respiratory symptoms and cardiovascular disease (CVD) (Bell et al., 2011). Acute asthma events, heart rate variability (Wu et al., 2010), birth weight (Rich et al., 2015), and bronchial hyperreactivity (Wong et al., 1998) have been studied as well. Most studies relate time periods of regulatory actions or changes in emissions to the health outcome of interest while controlling for environmental or population-related confounders, including meteorology, time trends, location, and age. Researchers then apply these models to a baseline pollution level (either from before/after the intervention of interest or a counterfactual) to estimate the relative risk of the response of interest to the regulatory

action. Some extend this to estimate the number of outcomes avoided due the regulation being adopted.

Zigler and Dominici (2014) argued that many studies (particularly those that compare periods before and after a regulatory action) do not account for all the factors that may impact a changing health response, and fail to address the uncertainties in the results relating to such an assumption. Cox (2013) discussed appropriate statistical methods for determining causality independent of a researcher's personal biases, such as Granger causality, conditional independence, and counterfactual causality tests.

2.3 Scope of this Work

This review assimilates work to date in the emerging field of accountability research. The studies are limited to those that assess air quality regulations in terms of at least one of the links in the accountability chain, and the focus is on studies that assess past regulatory action—as opposed to work that projects the effects of future regulations—or related events, such as a recession. A previous review (Bell et al., 2011) discussed publications focused specifically on the health endpoint. The current review extends their analysis by summarizing work that assesses other steps in the accountability chain, adding studies that have been published since 2011, and discussing future directions and challenges of accountability research.

In the next section, "Accountability Study Approaches", we present popular approaches taken by previous studies to investigate regulatory impacts on links in the accountability chain. Studies are grouped based on their endpoint of interest; for example, some studies are interested only in the regulatory impact on air quality, while others are interested in the impact on health effects. The following section, "Accountability Case Studies", reviews specific regulatory actions and the related works. The focus here is on discussing the ability of research methods to control for confounding factors. The next section reviews three previous commentaries on accountability studies. In the final two sections, we highlight challenges and future directions in accountability research, and conclude by summarizing the work and commenting on which aspects of accountability research deserve the most attention moving forward.

2.4 Accountability Study Approaches

Most accountability studies chose as an endpoint one of the links in the Accountability Chain: emissions, air quality, or health outcomes, though some addressed each step (Table 2-1 and Table 2-2). The endpoint of interest (i.e. emissions, air quality, or health outcomes) is associated with the depth of analysis at each link in the accountability chain. Columns in Tables 1-1 and 1-2 describe the types of analysis at each of the links in the Accountability Chain for each study. While health and air quality are the most common endpoints assessed (indeed, each of the studies that assessed health as an endpoint made an attempt to characterize the change in air quality associated with the regulatory action of interest), a few studies focused on the effectiveness of regulatory actions at reducing emissions. Implementation time scales investigated by the studies in this review fall into two categories: those implemented over a relatively short-term time frame (i.e., associated with hosting special events, such as the Olympics), and those implemented gradually over a long time scales (on the order of years) (Table 2-1 and Table 2-2). The studies focused on varied geographical scales, from single cities to global. The following subsections describe the approach of studies that define each of the three endpoints, and precede examples of specific case studies.

Table 2-1. Accountability studies with emissions and air quality endpoints.

Author	Year	Policy/event	Geographic scale	Time scale	Emissions	Air quality
Rao and Zurbenko	1994	All regulations, 1983-1991	Single monitor, New York State	Long		Ambient, meteorological detrending
Kuebler et al.	2001	All regulations, 1985-1998	National: Switzerland	Long	Changes in emissions by source over time	Ambient, meteorological detrending
Greenstone	2004	1970 Clean Air Act Amendments	National: U.S.	Long		Ambient, change in SO_2 concentrations in designated attainment vs. nonattainment regions
Camalier et al.	2007	All regulations, 1997-2005	Region: Eastern U.S.	Long		Ambient, meteorological detrending
Gégo et al.	2007	NO_x SIP Call	Region: Northeastern U.S.	Long	Changes in point-source emissions over time	Ambient, meteorological detrending (including wind direction analysis)
Gégo et al.	2008	NO_x SIP Call	Region: Northeastern U.S.	Long	Changes in point-source emissions over time, estimate of counterfactual	Modeled, comparing actual and counterfactual using single meteorological year
Godowitch et al.	2008	NO_x SIP Call	Region: Northeastern U.S.	Long	Changes in point-source emissions over time, estimate of counterfactual	Modeled dynamic analysis, comparing change due to emissions vs. meteorology
Ban-Weiss et al.	2008	Mobile regulations, 1997-2006	Local: tunnel	Long	Measured emissions factors, before/after comparison	
Atkinson et al.	2009	Congestion charging scheme, 2003	City: London	Short		Ambient, before/after comparison
Goodman et al.	2009	Coal sale bans	National: Ireland	Long		Ambient, before/after comparison
W Wang et al.	2009	Industrial activity reduced for 2008 Olympics in Beijing	City: Beijing	Short		Ambient, before/after comparison, back-trajectory analysis, correlations with meteorology
M Wang et al.	2009	Reduced industrial activity for 2008 Olympics in Beijing	City: Beijing	Short		Ambient, before/after comparison

Table 2-1. (Continued)

Author	Year	Policy/event	Geographic scale	Time scale	Emissions	Air quality
Wang and Xie	2009	Reduced industrial activity for 2008 Olympics in Beijing	City: Beijing	Short	Modeled, before/after comparison	Modeled, before/after comparison
X Wang et al.	2009	Reduced industrial activity for 2008 Olympics in Beijing	City: Beijing	Short	Measured emissions factors, before/after comparison	Ambient, before/after comparison
Y Wang et al.	2009	Reduced industrial activity for 2008 Olympics in Beijing	City: Beijing	Short	Modeled, before/after comparison	Modeled, before/after comparison
Kean et al.	2009	Mobile regulations, 1999-2006	Local: tunnel	Long	Measured emissions factors, before/after comparison	
Godowitch et al.	2010	Mobile regulations, 2002-2006	Region: Eastern U.S.	Long	Changes in emissions by source over time	Modeled dynamic analysis, comparing change due to emissions vs. meteorology
Dallman and Harley	2010	Mobile regulations, 1996-2006	National: U.S.	Long	Changes in modeled mobile emissions over time	
B Wang et al.	2010	Reduced industrial activity for 2008 Olympics in Beijing	City: Beijing	Short	Measured emissions factors, before/after comparison	Ambient, before/after comparison, source apportionment
S Wang	2010	Reduced industrial activity for 2008 Olympics in Beijing	City: Beijing	Short	Modeled, before/after comparison	Ambient and modeled, before/after comparison
Lin et al.	2010	All regulations, 2004-2008	Region: Eastern China	Long	Changes over time estimated by satellite	Ambient, change over time (satellite measurements)
Lamsal et al.	2011	All regulations, 2003-2009	Global	Long	Changes over time estimated by satellite	
Colette	2011	All regulations, 1998-2007	Continent: Europe	Long	Changes in emissions by source over time	Modeled dynamic analysis, comparing change due to emissions vs. meteorology
Butler et al.	2011	All regulations, 1997-2008	Region: Eastern U.S.	Long	Changes in point-source emissions over time	Ambient, change over time and sensitivity to emissions
Xing et al.	2011	Reduced industrial activity for 2008 Olympics in Beijing	City: Beijing	Short	Counterfactual emissions	Modeled dynamic analysis, comparing change due to emissions vs. meteorology

Table 2-1. (Continued)

Author	Year	Policy/event	Geographic scale	Time scale	Emissions	Air quality
Kelly et al.	2011	Congestion charging scheme introduced in 2003	City: London	Short	Modeled spatial	Ambient and modeled, before/after and control location comparison, source apportionment
Hao et al.	2011	Shanghai World expo 2010	City: Shanghai	Short		Satellite-measured, before/after comparison
Harrington et al.	2012	1990 CAAA	Region: Eastern U.S.	Long	Counterfactual	Ambient, modeled counterfactual
McDonald et al.	2012	Mobile regulations, 1990-2010	Multi-city: U.S.	Long	Changes over time	
Shreifels et al.	2012	SO2 reduction goals in the 10th and 11th Fiver-year Plans	National: China	Long	Changes over time	
Morgentsern et al.	2012	Title IV Phase 2 of the 1990 CAAA	Region: Eastern U.S.	Long	Change over time and counterfactual	Ambient, trends over time and counterfactual
Pegues et al.	2012	1997 8-hour ozone State Implementation Plans	National: U.S.	Long	Modeled, change over time	Ambient, change over time and location
Liu et al.	2012	Reduced industrial activity for 2008 Olympics in Beijing	City: Beijing	Short		Ambient and satellite-measured, statistically-modeled changes attributable to meteorology vs. emissions
He et al.	2013	All regulations, 1997-2011	Region: Eastern U.S.	Long	Changes in point-source emissions over time	Ambient, change over time
Yuan et al.	2013	All regulations, 1998-2008	City: Hong Kong	Long	Changes over time	Ambient, change over time, source apportionment, and sensitivity to emissions
McDonald et al.	2013	Mobile regulations, 1990-2010	Multi-city: U.S.	Long	Changes over time	
Kuwayama et al.	2013	Adoption of Clean Diesel Technology at a Major Shipping Port	Local: port	Short	Measured emissions factors, before/after comparison	Ambient, before/after comparison, source apportionment
Huang et al.	2013	Shanghai World expo 2010	City: Shanghai	Short		Ambient, before/after comparison, back trajectory analysis

Table 2-1. (Continued)

Author	Year	Policy/event	Geographic scale	Time scale	Emissions	Air quality
Lin et al	2013	Shanghai World expo 2010	City: Shanghai	Short		Ambient, before/after comparison, back trajectory analysis
Lurmann et al.	2014	All regulations, 1994-2011	Region: Southern California	Long	Changes over time	Ambient, change over time
Sickles II and Shadwick	2014	All regulations, 1990-2009	Region: Eastern U.S.	Long	Changes over time	Ambient, change over time
Wang et al.	2014	SO_2 and NO_x control policies, 2006-2015	Region: Eastern China	Long	Counterfactual	Modeled counterfactual
Vijayaraghavan et al.	2014	Mobile regulations, 1995-2010	City: Atlanta, GA	Long	Modeled actual	Summertime ozone (detrended)
Harley	2014	California Drayage Truck Regulation	Local: Port of Oakland	Long	Measured emissions factors, change over time	
Simon et al.	2015	All regulations, 1998-2013	National: U.S.	Long	Change over time	Ambient, trends over time
Liu et al.	2015	Reduced industrial activity for 2008 Olympics in Beijing	City: Beijing	Short		Ambient, neural network analysis and 1-D box model
Russell et al.	2016	Closure of three power plants	City: Pittsburgh, PA	Short	Fuel use used as a proxy	Ambient, before/after comparison, statistically-modeled change due to emissions
Van der A. et al.	2016	All regulations, 2005-2015	National: China	Long	Changes over time estimated by satellite	
Daskalakis et al.	2016	All regulations, 1980-2010	Global	Long	Counterfactual	Modeled counterfactual
Henneman et al.	n.d.	All regulations, 1993-2013	City: Atlanta, GA	Long	Counterfactual	Ambient, before/after comparison, source apportionment, and counterfactual

Table 2-2. Accountability studies with health endpoints.

Author	Year	Policy/event	Geographic scale	Time scale	Emissions	Air quality	Dose/ exposure	Health response
Pope	1989	Brief closing of steel mill, 1987-1988	City: Utah Valley	Short	Discussion of local contribution	Ambient, before/after comparison	Monthly mean and lagged mean	Respiratory illness, changes in health response rates before/after
Peters et al.	1996	1990 reduction of sulfur in fuel oil	City: Hong Cong	Short		Ambient, before/after comparison		Household survey, changes in health response rates before/after
EPA Section 812 report	1997	1970 Clean Air Act Amendments	National: U.S.	Long	Counterfactual	Ambient, change over time and modeled counterfactual	Varied across pollutants	Multiple outcomes, change from observed to counterfactual
Wong et al.	1998	1990 reduction of sulfur in fuel oil	City: Hong Cong	Short		Ambient, before/after and more/less polluted district comparisons		Bronchial hyperreactivity, changes in health response rates before/after and in more/less polluted districts
Friedman et al.	2001	Transportation limits for 1996 Olympics	City: Atlanta	Short	Before/after comparison	Ambient, before/after comparison	Daily, 2- and 3-day lag city average	Asthma events in children, changes in health response rates before/after
Clancy et al.	2002	1990 coal sale ban	City: Dublin	Short		Ambient, before/after comparison		Mortality, changes in health response rates before/after
Hedley et al.	2002	1990 reduction of sulfur in fuel oil	City: Hong Cong	Short		Ambient, before/after and more/less polluted district comparisons		Mortality, changes in health response rates before/after
Chay and Greenstone	2003	1981-1982 recession	National: U.S.	Short		Ambient, before/after comparison	Annual county average	Infant mortality, changes in health response rates before/after
Chay et al.	2003	1970 Clean Air Act Amendments	National: U.S.	Long		Ambient annual average by county	Annual county average	Adult mortality, changes in health response rates before/after
Laden et al.	2006	Reductions in PM2.5, 1979-1998	Multi-city: U.S	Long		Ambient, trends over time	Four annual averaging methods	Mortality, changes in health response rates with changes in concentration
Lee et al.	2007	Transportation limits for 2002 Asian Games	City: Busan, Korea	Short		Ambient, before/after comparison	Daily city average	Asthma events in children, changes in health response rates before/after
Pope et al.	2009	Reductions in PM2.5, 1970s-2000s	National: U.S.	Long		Ambient, trends over time	Annual mean	Mortality, changes in health response rates with changes in concentration
Peters et al.	2009	Rapid industrialization of a city, 1990-2002	City: Erfuhrt, Germany	Long	Change over time	Ambient, trends over time	Daily, single monitor lagged up to five days	Risk of death, change over time

Table 2-2. (Continued)

Author	Year	Policy/event	Geographic scale	Time scale	Emissions	Air quality	Dose/ exposure	Health response
Hou et al.	2010	Reduced industrial activity for 2008 Olympics in Beijing	City: Beijing	Short	Before/after comparison	Ambient, before/after comparison	Population-weighted exposure	Mortality and morbidity, changes in health response rates before/after
Li et al.	2010	Reduced industrial activity for 2008 Olympics in Beijing	City: Beijing	Short		Ambient, before/after comparison	Daily city average	Asthma events, changes in health response rates before/after
Wu et al.	2010	Reduced industrial activity for 2008 Olympics in Beijing	City: Beijing	Short		Personal exposure, before/after comparison	Personal exposure monitoring	Association of heart rate variability with PM2.5 exposure
Peel et al.	2010	Transportation limits for 1996 Olympics	City: Atlanta	Short	Traffic counts	Ambient, change over time and location		Relative risk for respiratory hospital visits compared with baseline period
Zigler et al.	2012	1990 CAAA – NAAQS attainment designation	Regional: Western U.S.	Long		Ambient, counterfactual 3-year average	3-year average county average	Mortality, change from observed to counterfactual
Rich et al.	2012	Reduced industrial activity for 2008 Olympics in Beijing	City: Beijing	Short		Ambient, before/after comparison	Daily, single monitor	Multiple outcomes, changes in response rates before/after and in various districts
Wong et al.	2012	1990 reduction of sulfur in fuel oil	City: Hong Kong	Short		Ambient, change over time and location	Daily, 1- and 2-day lag city average	Mortality, changes in health response rates before/after
Lin et al.	2013	NOx Budget Trading Program	Regional: New York State	Long		Ambient, before/after and more/less polluted district comparisons	Kriged 3-day moving average	Multiple health outcomes, changes in health response rates before/after and in various districts
Dockery et al.	2013	Coal sale bans	National: Ireland	Long		Ambient, change over time and location		Mortality, changes in health response rates before/after
Rich et al.	2015	Reduced industrial activity for 2008 Olympics in Beijing	City: Beijing	Short		Ambient, before/after comparison	Daily, single monitor	Change in birth weight associated with improved air quality during specific months in pregnancy
Su et al.	2015	Reduced industrial activity for 2008 Olympics in Beijing	City: Beijing	Short		Ambient, before/after comparison and 72-hr back trajectory analysis	Daily, single monitor, 0-4 day lag	Cardiovascular disease mortality, changes in health response rates before/after
Zigler et al.	2016	1990 CAAA – various regulations	National: U. S.	Long		Causally modeled		Multiple health outcomes, change from observed to counterfactual

2.4.1 Emissions Endpoint Studies

In the United States, multiple studies have assessed the impacts of new regulations and a changing fleet on mobile emissions, and have found significant reductions in emissions of multiple pollutants since the 1990s (Ban-Weiss et al., 2008; Dallmann and Harley, 2010; Harley, 2014; Kean et al., 2009; Kuwayama et al., 2013; McDonald et al., 2013, 2012). The National Emissions Inventory (NEI), assembled by the EPA every three years, is one example of an emissions endpoint study (https://www.epa.gov/air-emissions-inventories/national-emissions-inventory). The NEI seeks to quantify emissions from all sources of criteria air pollutants, criteria pollutant precursors, and hazardous air pollutants. Development of the NEI has employed a variety of methods (primarily bottom-up) depending on the source to assimilate emissions estimates and depends on data from each of the states. An example of bottom-up emissions estimates is large point sources, which are required to measure their stack emissions and report these to the EPA (U.S. EPA, 2016).

One type of top-down emissions estimate is using satellite measurements of column concentrations of air pollutants. The raw data from satellites must be treated by models that make numerous assumptions, such as the shape of the pollutant's vertical profile and the relationship between the actual measurement to the pollutant of interest (Streets et al., 2013). Satellite data have the benefits of being available with relatively small time lag than other emissions assimilating techniques and covering a wide spatial scale. Limitations include measurement drift, interference from clouds and surfaces with high albedo, and temporal coverage, which depends on each satellite's orbit. Streets et al. (2013) review multiple studies that have used satellite measurements to track changes in emissions at various spatial and temporal scales (Lamsal et al., 2011; Lin et al., 2010; van der A et al., 2016). Lamsal et al. (2010) compare satellite inventories to bottom-up inventories from different world regions, and finds an agreement of 6.0% globally (where 6.0% represents the relevant fraction of the bottom-up inventory) between the two.

The majority of emissions endpoint accountability studies compare emissions before/during/after an intervention. Studies that assess regulations implemented over many years (e.g. the 1990 CAAA), must account for longer-term trends than those impacted by controls, such as trends in fuel use, miles traveled, etc. Accountability studies focused on events that occur on short time scales (e.g. policies for the Atlanta and Beijing Olympics,

which lasted a few weeks) ignore long-term emissions trends (Kuwayama et al., 2013; Li et al., 2010).

Others have used measurements at high resolution to estimate changes in emissions factors in real-world driving situations (Ban-Weiss et al., 2008; Kean et al., 2009; Kuwayama et al., 2013). Estimates have been made both for fleet averages (Ban-Weiss et al., 2008; Kean et al., 2009) and specific vehicles matched by registration number (Harley, 2014). Alternatively, researchers have applied source apportionment techniques to ambient measurements to estimate emission changes (Huang et al., 2012; Kuwayama et al., 2013; B. Wang et al., 2010). In China, Schreifels et al. (2012) compared SO_2 emissions reductions under different periods of environmental policies (Five-year Plans), and found a 28% increase in emissions during the 10th Five Year Plan (2001-2005) and a 14% decrease during the 11th Five Year Plan (2006-2010). The government's goals were 10% reductions in both of the periods.

For emissions estimates made using emissions factors and activity level, it is straightforward to estimate a counterfactual realization of emissions that assumes no added controls under the regulation of emissions and the same activity level. This approach is taken in a number of studies that focus on the air pollution endpoint (Daskalakis et al., 2016; Gégo et al., 2008; Morgenstern et al., 2012).

2.4.2 Air Pollution Endpoint Studies

A second group of studies focus on the impacts of emissions regulations on air quality. Researchers place as a high priority the comparison of the magnitudes of effects of changing meteorology versus changing emissions on measured or modeled concentrations, and studies with this endpoint usually assess long-term interventions. Two main approaches are used to address the linkage between emissions and air quality: empirical approaches, including statistical modeling, often with meteorological detrending (Butler et al., 2011; Camalier et al., 2007; Gégo et al., 2007; Henneman et al., n.d., 2015; Kuebler et al., 2001; Liu et al., 2012; Rao and Zurbenko, 1994) and chemical transport modeled dynamic analysis (Colette et al., 2011; Gégo et al., 2008; Godowitch et al., 2010, 2008; Xing et al., 2011). A third approach tests directly whether designations of non-attainment, which required further actions by local governments, has led to the area

achieving its goals of improved air quality (Greenstone, 2004; Pegues et al., 2012; Zigler et al., 2016).

Empirical assessments of how the atmosphere responds to controls typically rely on developing statistical relationships of long-term ambient observations with emissions. These relationships, however, are confounded by meteorological variations, particularly in shorter-term studies. Identifying atmospheric responses over shorter periods–e.g., less than multiple years–can be difficult, and over periods of less than one year, meteorological variations dominate (Camalier et al., 2007; Rao and Zurbenko, 1994). Two events and associated accountability studies (described in more detail below) that note the difficulty in accounting for meteorological variability are the Atlanta Olympics Studies (Friedman et al., 2001; Peel et al., 2010) and the Shanghai World Expo (Y. Lin et al., 2013).

Air quality models provide a second, complimentary approach to link emissions to air quality changes. These models directly provide emissions-air quality relationships by describing and following emissions, transport, transformation, and fate in the atmosphere. However, they are subject to uncertainties in inputs and model parameterizations (Napelenok et al., 2011). In dynamic analysis, researchers model two (or more) time periods and assess the ability of the model to capture observed trends across the time span of interest, with a particular interest in controlling for meteorological variability (Colette et al., 2011; Gégo et al., 2008; Godowitch et al., 2010, 2008; Kang et al., 2013; Napelenok et al., 2011; Pierce et al., 2010; Xing et al., 2011). Some studies then repeat the modeling over the same domain with switched meteorological and emissions fields (Foley et al., 2015a, 2015b; Godowitch et al., 2008). The differences in modeled concentrations between model runs with the same meteorological inputs and different emissions represent the change attributable to emissions, and the differences between runs with different meteorology and the same emissions are attributable to meteorological changes.

Other air pollution studies seek to link changes in air quality with changes in emissions using statistical modeling techniques. These range from linear models (Harrington et al., 2012; Henneman et al., n.d.) to principal components analysis (PCA)-related and other source apportionment techniques (Balachandran et al., 2012; Blanchard et al., 2012; Hopke, 2016). Each source apportionment technique carries with it

uncertainties, and there have been recent efforts in this domain to combine results from multiple methods (Balachandran et al., 2012; Hopke, 2016).

In general, studies focused on estimating the change in air pollution due to emissions changes exert more effort explicitly accounting for differing meteorological conditions, while those focused on health do not when assessing changing air quality. However, most health studies do account for meteorology in their health models. By more fully accounting for regional meteorological patterns around the time of the 1996 Olympics in Atlanta, Peel et al. (2010) were unable to associate improved air quality with emissions reductions, a link that was found earlier by Friedman et al. (2001).

2.4.3 Health Endpoint Studies

Health outcome studies attempt to answer the question "what change in outcome response (if any) can be associated with the intervention of interest?" Such studies begin with an assessment of the intervention, which usually involves a comparison of air quality and health outcomes before and after the intervention using descriptive statistics (e.g. Dockery et al., 2013; Peters et al., 1996). Some studies couple this with a discussion of the meteorological conditions before, during, and after the intervention. Most employ a regression analysis to link changes in a measure of public health (e.g. morbidity or mortality) to changes in air quality, either by estimating outcome rates before/after an intervention or using direct associations of the outcome with a measure of air quality. Studies that seek to estimate the direct response of the health outcome to air pollution include a measure of the exposure of air pollution (e.g. annual average, multiple day lag, etc.) (e.g., Laden et al., 2006; Lee et al., 2007; Peters et al., 2009; Pope et al., 2009; Wu et al., 2010). Many studies have linked increased risk of disease and death, however, there is uncertainty in the shape of the concentration-response (C-R) function; there is evidence that the C-R function is supralinear, which provides evidence that potential benefits of reducing air pollution are greatest in relatively clean areas (Pope et al., 2015). Other studies do not estimate air pollution exposure, and instead either use factor variables to address changing conditions before and after the intervention or train different models for different time periods (e.g., Clancy et al., 2002; Peel et al., 2010). Regressions include many independent variables to reduce confounding that are chosen based on the data available, researchers' initial understanding of the problem, and the model selection process, which

usually involves removing covariates with high standard errors one by one (Pope et al., 2012). Results are reported in the form of change in health response rates or excess risk from air pollution attributable to the intervention.

Many accountability studies that assess health impacts do not fully address the relationships between regulations, emissions, air quality, and dose, a recommendation made by the HEI (Health Effects Institute, 2003). Often (particularly in short-term studies), the baseline scenario is taken as the time before the intervention, an assumption that fails to address changes in confounding variables that may occur simultaneously with the intervention (Zigler and Dominici, 2014). In long-term studies, confounders that impact health and are correlated with changing air pollution concentrations (such as access to health care) make finding statistical associations between life expectancy and air pollution policies difficult (Wong et al., 2012). Further, health studies use meteorological metrics to control health response models, but do not assess air pollution changes while controlling for meteorology or locations surrounding study area. Many studies of short-term events do not investigate changes in other emissions at the same time as the intervention.

Recent research has discussed the issues with confounding and scientist biases that constitute an issue with the approach of linking health outcomes with air pollution concentrations and estimating a counterfactual (Cox and Popken, 2015; Pope et al., 2012; Zigler and Dominici, 2014; Zigler et al., 2016). Cox and Popken (2015) and Zigler and Dominici (2016) argue for and apply more advanced statistical tests and methods that go further to assess the causality (instead of the correlation) of air pollution outcomes. These authors argue that the typical approach, which uses models that associate observed concentrations with health outcomes, is inappropriate for the question of interest. They posit that a more appropriate question, 'What is the relationship between a specific regulation and health?' better serves the goal of accountability research. In order to answer this question, Cox and Popken suggest using such methods as conditional independence and Granger Causality tests, which are statistical methods that better account for alternative explanations for correlations between, for example, air pollution and health. Zigler and Dominici (2016) use two statistical methods—principal stratification and causal mediation analysis—to test for causal linkages between long-term air pollution interventions, air quality, and public health.

2.4.4 Comparison with Pre-Regulation Impact Assessments

The United States EPA publishes detailed analyses of the anticipated benefits of a regulation before it is promulgated in regulatory impact assessments (RIAs) (e.g. U.S. Environmental Protection Agency, 2005a, 2005b, 1999, 1997). For air quality regulations, RIAs estimate effects of the regulation on emissions, air quality, and human/ecosystem health. Further, the EPA estimates an expected cost of the regulation based on the controls or other changes needed to meet emissions targets.

The 2003 HEI report recommends using outcomes assessed in RIAs as starting points for accountability assessments in order to evaluate the conclusions of the RIA. Harrington et al. (2012) compared the results of their accountability assessment of the Clean Air Interstate Rule (CAIR) with projected impacts reported in the RIA from the same rule (U.S. EPA, 2005b). They note that direct comparison is difficult due to differences in the methods used and context of the two studies. The authors compare both the final results (changes in $PM_{2.5}$ concentrations in the Eastern United States attributable to the 1990 Clean Air Act Amendments) and the sensitivities of $PM_{2.5}$ to utility emissions. They find that their post-intervention estimates of air quality changes generally align. Given the current trend of assessing the accountability of longer-term regulatory actions (many of which were accompanied by an RIA), the comparison of accountability analyses with impact estimates prior to the regulation is an important step to assess and improve the regulatory process.

2.5 Accountability Case Studies

This sections reviews policy and intervention scenarios that have attracted interest in accountability research. Many of these have been investigated multiple times, some with updated methods each time and some with updated data (e.g. emissions, air pollution measurements, or health outcomes data).

2.5.1 1986-87 Utah Valley Steel Mill Closing

The Geneva steel mill, a major source of air pollution in Utah Valley, closed for a period of one year spanning 1986-87 due to a labor dispute (Pope, 1989). While the closing was not due to a regulation, the analysis conducted on the resulting health impacts is of interest in accountability studies, as similar methods may be used to assess potential health impacts resulting from regulations leading to emissions changes. Pope related ambient

measurements of particulate matter with a diameter less than or equal to 10 μm (PM_{10}) and respiratory-related hospital admission using a multiple regression model. They used hospital admissions outside the county as a control, and further controlled for temperature cofounders. The researchers observed nearly doubled PM_{10} concentrations in the valley when the mill was open compared to when it was closed, and found associations between multiple health outcomes and PM_{10} levels across the time periods before, during, and after the intervention.

A goal of this study was to determine whether negative health outcomes could be associated with ambient air pollution concentrations over a period of such a large change. The author discussed other sources of emissions in the valley (noting that they were small in comparison to the mill), compared air quality before, during, and after the intervention, used different estimates of dose in the health outcomes regression models, and controlled for various confounders. In total, results from 18 models with various combinations of health outcomes and confounders were reported.

The results of this laid groundwork for future accountability studies of short-term interventions. Pope showed that the association of PM_{10} with various health outcomes was robust to a number of different models across a period of changing emissions, and used similar regressions with health outcomes at nearby hospitals to control for coincidental decreases in influenza when the plant was closed. The author noted that PM_{10} may serve as a surrogate measure for other pollutants, but did not find evidence to suggest that PM_{10} was serving as a surrogate for ozone.

2.5.2 1990 Dublin Coal Ban

In response to poor air quality in Dublin attributed to the use of coal for heating during the 1980s, the government banned coal sales in the city beginning on 1 September, 1990. The subsequent improvement in air quality provided an opportunity to assess the impacts of the ban on public health outcomes. Clancy et al. (2002) compared seasonal black smoke (BS), sulfur dioxide (SO_2), temperature, relative humidity, and non-trauma death rates for six years before the intervention and six years after. Death rates were estimated and adjusted in order to account for other changes happening in Dublin parallel with the ban, including meteorology, changing age distribution, and respiratory disease epidemics (influenza). Other changes that could not be captured were accounted for by adjusting death

rates by age-standardized death rates for the rest of Ireland, excluding Dublin. The authors found that BS concentrations decreased 70% after the ban and SO_2 levels were reduced 34%. The ban coincided with a reduction in respiratory and cardiovascular-related death rates similar in magnitude to those found by Pope (1989). They concluded that, even given decreasing death rates throughout Ireland, the coal ban led to substantial improvements in air quality and subsequent reductions in negative health outcomes.

Dockery and colleagues revisited the Dublin coal ban in later studies, along with similar bans in other cities in Ireland that occurred in 1995, 1998, and 2000 (Dockery et al., 2013; Goodman et al., 2009). In their reassessment, which included a detailed sensitivity analysis and a more comprehensive approach for correcting for long-term background trends unassociated with each ban and other secular effects, they found only a decrease in respiratory mortality rates in Dublin after the 1990 ban, and no significant reductions at other locations after subsequent bans. The interrupted time-series study design used here, in which the health model included a binary variable that indicating controlled/vs. not controlled periods, is a direct accountability technique (discussed in more detail below). The lack of association in the follow-up analysis was linked to better accounting for background health trends in other parts of the country unaffected by the ban (Dockery et al., 2013).

One drawback of the Dockery et al. (2013) study was the lack of measurement data in the control cities (i.e., those that did not implement a coal ban). The authors could draw conclusions on changes in air quality in districts that had the ban, but could not draw conclusions on how changes in these areas were confounded by similar changes seen throughout the country. Taken together, the two studies show how confounders—concurrent changes in societal factors in this case—can obscure results obtained in accountability research. In the later study, more data on both spatial and temporal scales increased the number of viable analyses available to researchers and increased the ability to draw sound conclusions regarding the effectiveness of control programs.

2.5.3 1990 Hong Kong Sulfur Ban

In 1990, Hong Kong began requiring fuel oil (used for power generation and transportation) sold in the city to be 0.5% by weight sulfur or less. Multiple studies have assessed the impacts of the reduction on air pollution and public health coinciding with the

ban (Hedley et al., 2002; Liu et al., 2015; Peters et al., 1996; W. Wang et al., 2009; Wong et al., 2012, 1998). Peters et al. (1996) compared different areas throughout Hong Kong, and found that after the ban, more polluted areas experienced large reductions in SO_2 and sulfate concentrations, and found subsequent associations with reduced respiratory symptoms in children. Peters et al. used multivariate logistic regression to estimate risk of respiratory problems related to a number of factors, including residence district, school, factors representing pre-and post-intervention, and multiple confounding variables. In this study and those that followed, researchers used results from surveys of children in primary school (Wong et al., 2012, 1998).

Hedley et al. (2002) assessed the change in mortality rates associated with the ban using ambient concentrations and death records from all causes. The researchers used a Poisson regression model that accounted for seasonal fluctuations, temperature and relative humidity changes to estimate the relative change in average annual numbers of deaths in the five-year periods before and after the intervention. They found a reduction in seasonal deaths in the first 12 months after the intervention, and a return to the expected death rate from the 3rd year to the fifth. The authors concluded that the coal ban did have an effect on air quality and public health outcomes the first year, and the subsequent return to the expected rate was attributed to a delayed death phenomenon. They discussed how secular changes in death rates and other factors may have had an effect on death rates, but did not adjust the results using a control population.

In the most recent follow-up, Wong et al. (2012) used long-term—5 years on either side of the ban—records of air pollution and mortality to investigate the effect of the intervention on lifespan. The study met difficulties, including the inability to separate health effects of different pollutants (nickel, vanadium, and SO_2), and large uncertainties caused by high spatial and temporal variability through the study period. In their review of the project, the HEI concluded that the inability to control for potential confounders that correlated with changes in air quality means that there is little hope for estimating changes in life expectancy using daily time series models (Wong et al., 2012).

The varied results in the Hong Kong studies show the difficulty in estimating impacts of regulatory actions, even when large amounts of data are available. An important aspect of each these studies was the comparison between more-and-less polluted districts,

which showed that more polluted districts experiences larger improvements than less previously less polluted districts. These areas, however, are not perfect controls, since they have many differences other than their pollution levels that could impact health outcomes (such as access to health care) (Wong et al., HEI Commentary, 2012).

2.5.4 1996 Atlanta Olympics

The 1996 Summer Olympics were held in Atlanta, GA, during which a local road traffic congestion management strategy was implemented for 17 days that included banning private automobile traffic in the downtown area, new options for public transportation, and alternative work hours for businesses. Friedman and colleagues (2001) compared air quality (multiple pollutants), childhood asthma and non-asthma acute events, meteorology, and vehicular traffic for the period during the Olympics to the four-week periods prior and post. They found significant decreases in traffic counts along with improvements in both air quality and public health over the period of the Games, but did note that the cause of the change in air quality was likely linked to both decreased emissions and favorable meteorological conditions. They concluded that, even with the influence of ideal meteorological conditions, the policies related to the Games had an impact on air quality and public health in Atlanta.

Peel et al. (2010) revisited the same intervention using an expanded dataset of emergency department visits for 1995 through 2004, air quality and meteorological data from multiple monitoring sites in Atlanta and the surrounding counties, and traffic counts at 18 sites within the five counties of interest. Similar to Friedman et al., Peel et al. reported reduced concentrations of ozone and other air pollutants in Atlanta during the games, but further concluded that, in light of lower ozone levels across the Southeastern United States and meteorological conditions unfavorable to ozone production across the region, the main cause of improved air quality during the games was improved atmospheric conditions, and not a change in emissions. They did not find significant reductions in respiratory or cardiovascular emergency department visits in adults or children during the Games after adjusting for seasonal trends and other confounders, but the results were limited by a short time period of interest and low numbers of daily emergency department visits.

Peel et al.'s study was limited by the intervention's short duration, low numbers of emergency department visits, lack of control areas, and a difficulty in isolating the impact

of the control from typical temporal patterns in health effects (Peel et al., HEI Commentary, 2010). The Atlanta Olympics studies provide examples of accountability studies that investigate an intervention from the perspectives of both the classic accountability paradigm and using the direct approach (i.e., using statistical methods that link the control period directly to the health outcome of interest).

2.5.5 2008 Beijing Olympics

As a condition for hosting the 2008 Olympics, the Chinese government agreed to reduce emissions in order to improve the air quality for both the Games and the ensuing Paralympic Games (Rich et al., 2012). This chain of events led to a number of studies assessing the effectiveness of the policies at reducing air pollution in the city and any associated health benefits (Hou et al., 2010; Li et al., 2010; Liu et al., 2015, 2012; Rich et al., 2015, 2012; Su et al., 2015; B. Wang et al., 2010; M. Wang et al., 2009; S. Wang et al., 2010; Wang and Xie, 2009; W. Wang et al., 2009; X. Wang et al., 2009; Y. Wang et al., 2009; Wu et al., 2010; Xing et al., 2011).

An advantage to the planned emissions reductions in Beijing was that researchers had the opportunity to plan air quality observational studies, recruit test subjects beforehand, and test health outcomes during the intervention period—a major shift compared to the historical approach to accountability research of natural experiments designed after an intervention. Researchers planned ambient measurement campaigns (B. Wang et al., 2010; M. Wang et al., 2009; X. Wang et al., 2009) and used measurements of personal exposure (Wu et al., 2010) and health markers (Rich et al., 2012) to assess impacts of emissions reduction. Using data on birth weights in the Beijing area, Rich et al. (2015) found that babies in their 8^{th} month of gestation during the Olympics were statistically significantly larger at birth than babies born at the same term the year before and after. Su et al. (2015) related CVD death counts to ambient pollution measurements while accounting for meteorology and air mass origin, and found a positive association. Hou et al. (2010) extended the analysis to an economic impact estimated based on value of statistical life and average outpatient costs in Beijing. All of the studies listed above found positive associations between reduced health outcomes and the time period of emissions reductions.

Xing et al. (2011) used a chemical transport model to estimate gas and particulate sensitivities to emissions and meteorology, and concluded that emissions controls during the period reduced the sensitivity of O_3 chemistry to VOC emissions in Beijing. They noted the limitations of assuming a linear system in their modeling, but asserted that the results show how reduced emissions controls have important effects on atmospheric conditions even in time periods of meteorology favorable to reduced pollution.

While most studies concluded that emissions controls led to improved air quality during the Olympics, Wang et al. (2009) used four day air plume back trajectories regressed against PM concentrations to conclude that 40% of the variability in ambient measurements was attributable to meteorology (compared to 16% attributable to emissions). The authors observed statistically significant changes in PM concentrations during the games compared to before and the Olympic games, but did not find statistically significant differences between source control and non-source control periods (controls were imposed a few weeks before the Games began). Wang et al. (2009) determined that wind direction from the cleaner territory to the northeast of the city during the Games created favorable conditions that reduced PM concentrations below what they would have been with just the controls.

The Beijing Olympics studies show the expanded types of studies available to scientists who can plan in advance to implement accountability studies. These include special ambient air pollution, personal exposure, and health marker measurement campaigns. Future accountability studies should take advantage of planned rules and plan ahead in a similar fashion, but should extend the planning to measurements of confounding factors (e.g. smoking habits, accessibility of care, etc.).

2.5.6 London Congestion Charging Scheme

On 17 February, 2003, London introduced a congestion charging scheme (CCS) in the center of the city, along with other regional traffic changes in an attempt to reduce traffic within the city during work hours on weekdays (Atkinson et al., 2009; Kelly et al., 2011). This intervention is an example of a policy not specifically designed to improve air quality, but one that may have potential benefits. Atkinson et al. (2009) used ambient concentrations measured at more than 40 monitoring cites in the greater London area to assess impacts of the CCS on air quality in the city. Using comparisons with geometric

mean background concentrations from monitors outside the city center, they were able to show evidence for decreased NO_x, O_3, PM_{10}, and CO within the CCS zone, however, given the concurrent traffic changes, the authors were not able to fully attribute causality to the CCS. Later, the same group used more detailed statistical analyses, including use of estimates of emissions changes attributable to the CCS, and again found small improvements in air quality within the CCS zone (Kelly et al., 2011). Further modeling using estimated emissions provided evidence that the changes were attributable to regional air quality regulations, such as the requirement for buses to install particulate filters.

These studies show the difficulty of quantifying the impacts of a single local regulatory action in the context of broader regional changes, variable meteorology, congruent regulations, and unanticipated institutional or behavioral changes that may be associated with a regulation. The results highlight the need for detailed estimates of uncertainties in measurements and modeling for each link in the accountability chain—the small improvements in air quality attributed to the CCS policy must be viewed with caution due to limited knowledge of the associated uncertainty (Kelly et al., HEI Commentary, 2011).

2.5.7 2010 World Expo in Shanghai

In 2010, Shanghai hosted the 6 month-long (May-October) World Expo, and implemented a number of emissions control measures to improve air quality (Hao et al., 2011; Huang et al., 2013, 2012; Y. Lin et al., 2013). The government imposed short-term limits on energy production, industries such as coking and cement-making, transportation, construction, and agricultural burning in the region.

Hao et al. (2011) compared satellite measurements of NO_2, aerosol optical thickness (AOT), and CO between the same period over the three previous years and six months after the Expo. They found reductions in NO_2, AOT, and CO of 8%, 14%, and 12% compared to before the Expo, and subsequent increases in NO_2 and AOT of 20% and 23% after the Expo. Huang et al. (2013) used high-resolution measurements of PM constituents to assess ambient pollution concentration trends during the Expo. They used back-trajectory analysis to identify important sources and regions (similar to the Beijing studies, they determined the Central Plains to be a major source). Assessment of the constituent make-up led to the conclusion that controls on stationary sources led to the reduction in

sulfate aerosols over the period of the Expo, but found significant increases in nitrate aerosol, which the authors attributed to increased traffic from visitors to the region during the Expo. This study discusses meteorological conditions, but does not adjust for meteorological fluctuations in their analysis.

Y. Lin et al. (2013) used measurements from 53 measurement sites in the Yangtze River Delta (YRD) region that includes Shanghai, and found that air quality was worse in May, June, and October, and that it was improved in the intermediary months. The researchers concluded the cleaner months aligned with wind patterns from the sea, which led to cleaner air near the coast and deteriorated air quality inland. They attributed lower SO_2 and PM_{10} concentrations during the Expo to limits imposed on power plants, industry, and construction, and higher NO_x, CO, and O_3 to increased vehicular traffic associated with visitors to the region for the Expo.

The study period for the Shanghai World expo had the benefit of being longer (6 months) than other short-term intervention studies (e.g. the Atlanta or Beijing Olympics). However, Lin et al.'s investigation into wind patterns led to the conclusion that meteorological variability and regional transport played an important role in the changes in air quality during this period. All three studies discussed here noted the role of an unanticipated outcome of an intervention—i.e. increased traffic due to visitors to the region. Their use of satellite and monitoring data, statistical descriptions and back-trajectory analysis provide an example of a multimedia approach to accountability research, which should be employed in further studies as available and applicable.

2.5.8 Port of Oakland Diesel Trucks Retrofit

Shipping ports are a major source of PM air pollution in coastal areas (Harley, 2014; Kuwayama et al., 2013). The state of California instituted controls on multiple sources associated with activities at ports, one of which was the Comprehensive Truck Management Plan, which began in October 2009. The program required that trucks meet emissions standards (requiring the use of a diesel particulate filter, or DPF) and register with the Port database by April 2010. Kuwayama et al. (2013) applied Positive Matrix Factorization (PMF), a source apportionment technique, to PM measurements near the port for weekdays between 8 March 2010 and 28 March 2010. They used the entire time period to train the PMF model, but only used a subset of days selected for wind speed and direction

coming from the Port in the analysis. The authors found that contributions of port truck traffic to PM components decreased by amounts ranging from 66% to 87%.

Harley (2014) measured concentrations at high time resolution at a major access road to the same port in 2011 and 2013. Emissions factors of passing trucks were estimated from the measurement data, and related to the passing vehicle using the registration database, which included information on model year and DPF retrofit status. The authors reported 99% coverage of DPF technologies, and reductions in fleet-averaged emissions factors of 76±22% and 53±8% of black carbon and NO_x, respectively over the period 2009-2013, which align with prior estimates by Kuwayama et al. (2013).

These studies show how before and after measurements can be made directly of emissions related to interventions, again exemplifying a benefit of planning an accountability study before a regulation goes in effect. Continuous emissions monitoring is available for large power plants in the United States (U.S. EPA, 2016), but availability is low for mobile sources.

2.5.9 Regulations under the 1990 United States Clean Air Act Amendments

Actions taken under the 1990 Clean Air Act Amendments (CAAA) presented researchers with an opportunity to extend accountability analysis from the short-term actions described above to actions taken incrementally over many years. Further, nation-wide ambient air quality networks made it possible to assess impacts over broad spatial scales. Between 1997 and 2011, estimated nationwide (United States) NO_x and CO emissions dropped by about 50% and 60%, respectively (He et al., 2013), and 95th percentile summertime ozone decreased over the period 1998-2013 by 1-2 ppb yr^{-1} (Simon et al., 2015).

Given the long time period and economic cost of the intervention, there have been many studies that focus on the emissions and air quality end points (Camalier et al., 2007; Gégo et al., 2008, 2007; Godowitch et al., 2010, 2008; Harrington et al., 2012; He et al., 2013; Lurmann et al., 2014; Sickles II and Shadwick, 2014; U.S. EPA, 2011). The lengthy data record of both air quality measurements and emissions estimates allows these researchers to use meteorological detrending and other statistical methods to assess the relative impacts of meteorology and emissions changes. Results agree across studies that the regulations have reduced air pollution concentrations.

The U.S. EPA undertook a detailed assessment of benefits and costs of the Clean Air Act as part of the requirements under Section 812 of the 1990 CAAA (U.S. EPA, 2011). They used modeled actual and estimated counterfactual emissions to drive a chemical transport model for two scenarios: "*with-CAAA*" and "*without-CAAA*", and estimated changes in concentrations, mortality, environmental impacts, costs, and benefits between the two. Estimated direct costs, most of which stemmed from control costs, totaled $53 billion (2006$) in 2010. The research used results of cohort air pollution studies and expert elicitations to decide on appropriate concentration-response functions for ozone and PM, and used population-weighted concentrations to estimate dose. These values were used with concentration-response functions and estimates of willingness to pay to estimate cases of 'excess premature mortality' per year and the sum of 'dollars per mortality avoided' (similar to 'value of a statistical life'–VSL). Estimated monetized benefits totaled $1.3 trillion (2006 $) in 2010. A detailed sensitivity uncertainty analysis to model inputs yielded variability in the results, but it was small compared to the difference between the estimates of total benefits and costs.

While most studies assess trends in emissions and air quality, a body of research uses a potential outcomes approach that estimates outcomes at links in the accountability chain for developing counterfactual emissions scenarios (Daskalakis et al., 2016; Harrington et al., 2012; Henneman et al., n.d.; U.S. EPA, 2011; Wang et al., 2014; Zigler et al., 2012). Morgenstern et al. (2012), for example, developed a statistical model that linked monthly $PM_{2.5}$ concentrations with power plant SO_2 emissions, and then used the model to estimate counterfactual concentrations. Henneman et al. (n.d.) extends this analysis to daily $PM_{2.5}$ and ozone. Daily results are important in this context for future use in acute health analyses.

Results for health studies are more varied (Cox and Popken, 2015; S. Lin et al., 2013; Pope et al., 2012; U.S. EPA, 1997b; Zigler et al., 2012). These studies have noted the difficulty in assessing long-term implementations of regulations due to the inability of statistical models to capture other changing factors that occur, such as shifting baseline health of a population or availability of care, statistical manipulation, or over interpretation of subgroup analysis. They employed alternative methods such as Granger causality tests (Cox and Popken, 2015) and principal stratification (Zigler et al., 2012).

Zigler et al. (2012) sought to attribute causality of improved public health to nonattainment designation for PM_{10} of counties in the Western United States under the 1990 CAAA. They used a statistical method called *principal stratification* to distinguish causal effects of a regulation on health that are associated with causal effects of the regulation on air quality verses those that are not. Their analysis covered 7 million Medicare enrollees in the Western United states, and finds that the effect of the nonattainment designation programs was 1.76 fewer deaths per 1000 Medicare beneficiaries.

Zigler et al. (2016) extended their previous efforts by assessing the designation of NAAQS nonattainment on air pollution in various locations using both principal stratification and causal mediation, a similar method that separates effects of an intervention into the 'natural direct' and 'natural indirect' effects. The authors employed a principal stratification method that matches areas based on their *propensity scores*—a statistical score that combines multiple confounders into one number—and uses locations not designated as nonattainment areas for the control. Their analysis shows that there is some interference between study and control locations. A limitation of this method is that, in order to compare multiple locations at once, propensity scores are developed across many varying conditions. First, for locations that are in nonattainment that have no matching areas that are in attainment (based on propensity scores), these locations cannot be used. Second, assigning a single score to a region based on a range of factors may dull the impact of one important factor. One important factor may influence air pollution or health outcomes more than many similar attributes.

The United States experienced widespread reductions in emissions and subsequent improvements in air quality between when the Amendments became law and the present. Collectively, the long-term Clean Air Act Amendment Studies represent an advance over the classical accountability studies that investigated short-term interventions. Results of studies focused on air quality find it straightforward to conclude that the legislation and regulations thereunder led to improvements. Studies that assessed health impacts found mixed results; some found large improvements, while others found small or no reductions in health effects. This collection of studies presents an argument for further development of statistical methods that are appropriate for assessing changes in health effects while

accounting for the multitude of confounders that change over long time periods on wide time scales.

2.6 Previous Commentaries on Accountability Studies

Key points from recent reviews of accountability analyses are discussed below, and a commentary on the direction of the field is provided in the following section.

2.6.1 Health Effects Institute, Communication 11, 2003

Health Effects Institute (HEI) Communication 11 in 2003 was intended to promote research aimed at refining methods for performing accountability analyses of current policies (Health Effects Institute, 2003). HEI made the case for thorough assessment of previous policies, namely, to inform future policy decision-making, and listed a number of opportunities for accountability studies. The work steps through the accountability chain and dissect confounders that influence the chain between the links. The authors listed potential issues that affect all accountability studies, and provided guidance on opportunities for future work. They presented broad recommendations in three areas: 1) development and testing of new study designs, 2) identifying targets for accountability research, and 3) development of systems to track health impacts of regulations. Specific challenges include: defining an appropriate baseline of pollutant levels, identifying the impact of mixtures of pollutants verses single pollutants, isolating the causal pathway between regulations and health outcomes independent of unanticipated changes in personal activities, behaviors, and other confounders. They caution that stakeholders often hold high expectations for accountability research, when the realities of the research may restrict the possibility of providing definitive evidence for efficacy of controls. Since the 2003 report was written, the HEI has funded a number of accountability studies (e.g., Dockery et al., 2013; Kelly et al., 2011; Peel et al., 2010; Peters et al., 2009; Wong et al., 2012; Zigler et al., 2016).

2.6.2 Review of Accountability Studies Assessing Health Endpoints, 2011

Bell et al. (2011) reviewed accountability studies that assessed health endpoints. They noted that the majority of the accountability studies published up to that point focused on short-term interventions at a local scale, and that extending the analysis to longer time scales and larger distance scales will take the development of more advanced statistical techniques. Increasing numbers of studies have improved their methodologies to take these

larger scales into account, such as those discussed above under the Clean Air Act throughout the United States. The authors identify several challenges for studies of this scale (both time and distance). Those discussed in depth include the choice of baseline pollution and health levels, transboundary pollutant transport from outside of the study area, and difficulties in exposure-response studies, such as air pollution modeling and extrapolation between populations and geographic areas.

Bell et al. note that it can be particularly difficult to determine baseline air pollution levels for secondary pollutants or ones that undergo nonlinear transformations in the atmosphere. They list a number of modeling approaches (e.g. source-receptor matrices and chemical transport models) that are used to estimate air pollution-emissions relationships, but caution that each of these modeling approaches carries with it limitations. The authors recommend further development and evaluation of these models and their extension to exposure estimates. The ability of a model to estimate absolute pollutant concentrations may not be as necessary as the model's ability to capture changes.

Exposure studies are subject to the same assumptions in accountability research as in more traditional epidemiological research. Examples of such assumptions includes the extrapolation of results across pollution levels (e.g. relating to the potential existence of a threshold pollutant level below which no adverse health effects are observed), populations, or time spans. Further, effects of multipollutant mixtures are difficult to assess because of limited scientific knowledge of interactions, correlation between pollutants, and chemical makeup of pollutants that are characterized together (e.g. $PM_{2.5}$ is composed of multiple chemical elements that differ between regions). They note the difficulty in assessing each and every health (and environmental) outcome of interest due to both data limitations and available statistical methods.

They conclude that, given the challenges inherent in accountability analysis, the most robust studies should include detailed assessments of uncertainties and present results from multiple models in order to provide cogent evidence that the policy did/did not impact links in the accountability chain.

2.6.3 American Journal of Epidemiology Point-Counterpoint, 2014

In 2014, the *American Journal of Epidemiology* (*AJE*) published a pair of commentaries in a point-counterpoint exchange that addressed the benefits and limitations

of potential-outcomes and prospective research designs (Hubbell and Greenbaum, 2014; Zigler and Dominici, 2014). Zigler and Dominici (ZD) argue that alternative-outcome research adds substance to the current body of air pollution policy research. They note that the majority of the research framed as accountability studies explore *indirect* relationships between emissions and health outcomes. These studies compare the time during/after the regulatory action (baseline) of interest to a relatable time before or after the action (Figure 2-1, right side). ZD argue that this ignores potential changes of the exposure-response relationship that may occur during the implementation of the regulatory action (as discussed in this article). Questions of *direct* accountability seek to attribute changes in health-response to well-defined regulatory actions while controlling for confounding by including comparable populations that are not subjected to the same regulatory action. Instead of quantifying effects of regulatory actions at each link in the chain, confounders are accounted for in statistical models that link the regulation of interest with the end point of interest (usually either air quality or health outcome). They state that direct studies are necessary to advance accountability research, and provide important evidence in the present regulatory climate.

In the counterpoint, Hubbell and Greenbaum note that the potential-outcomes framework offers an appealing alternative to traditional epidemiological methods that relate air pollution and health. However, they point out that the approach introduces additional factors (in particular, data availability and lack of randomness in the assignment of populations to control and study groups) that must be considered in the study design, and suggest that potential-outcomes studies would best be served by designing the studies before the intervention of interest in order to best account for confounders. Pope et al. (2012) acknowledge concerns—such as indiscriminant multiple testing of observational data sets, inadequate control of confounding variables, and statistical model manipulation—that relate to the application of observational studies. However, they note that the validity of each study of this type should be judged on its ability to account for confounding.

The direct accountability approach is somewhat at odds with the historical view of the accountability chain, in which each link should be evaluated for its relationship to the link before it. However, the difference in approaches may be subtler; when answering

questions of *direct* accountability, it remains important to account for potential confounders and control for changes over time that are simultaneous with regulatory actions. The confounders may be slightly different; for instance, if a direct accountability study assesses the impact of installing scrubbers on power plants on populations living near power plants, and an important health-related factor—e.g., smoking behavior—is comparable between the populations, smoking behavior is not a confounder (Zigler and Dominici, 2016).

2.7 Challenges and Future Directions

Two trends are apparent in the accountability assessments above. 1) Often, one study will assess a regulatory action and determine that the intervention led to a statistically significant change in the response of interest (Clancy et al., 2002; Friedman et al., 2001; Peters et al., 1996; Zigler et al., 2012). Later, using additional data, updated methods, and/or accounting for additional factors, those results are found to be less definitive and potentially invalid (Cox and Popken, 2015; Dockery et al., 2013; Peel et al., 2010; Wong et al., 2012). 2) Spatial and temporal scales of accountability studies have expanded, which has increased the number of potential confounders.

Both patters demonstrate the importance of appropriately accounting for confounding factors when developing models and understanding as much as practicable responses in each step of the accountability chain (Hubbell and Greenbaum, 2014). Examples of important confounders to consider include concurrent regulations, meteorological variability over time/space, interactions between the pollutant targeted by the regulation and other pollutants, and changes in overall health of a population of interest across time and space. In the classical accountability framework, information from these confounders is used to estimate baseline levels at each link in the accountability chain. Typically, researchers use a different type of model for each link, and assessments of both the model outputs (i.e., the counterfactual emissions, air quality, etc.) and the relationships in the model (e.g., the sensitivity of $PM_{2.5}$ concentrations to emissions) are necessary components of a full evaluations of baseline levels.

So-called direct accountability studies offer an alternative to the classical accountability framework. The direct framework has been used in other fields, and is becoming more popular in air pollution accountability. These methods should be

investigated further, and will likely complement the classical approach of investigating relationships in the accountability chain. Controlling for confounders is related to selecting an appropriate baseline, and statistical methods can be used to address confounding, but conventional methods may not be able to fully account for all confounders, and new statistical methods may be more appropriate for accountability studies (Cox and Popken, 2015; Pope et al., 2012; Zigler and Dominici, 2014).

When feasible, stakeholders should plan approaches to assess a regulation's efficacy *before* the regulation is implemented (as in the Beijing Olympics, where, for example, B. Wang et al. (2010) and X. Wang et al. (2009) began specialized air pollution measurement campaigns and Wu et al. (2010) measured continuous exposure and health markers in taxi drivers during the Olympic period). This will allow for study designs that more fully address confounding (instead of relying on the 'natural experiments' that have typically been used up to this point (Cox and Popken, 2015; Hubbell and Greenbaum, 2014). For example, studies can be designed to directly assess each link in the accountability chain (direct emissions measurements, expanded air quality observations in areas that are expected to most directly respond to the intervention, as well as control regions, and additional health monitoring).

Any changes attributable to regulations are relative to background pollution levels, which are a function of global emissions trends and climate change (Fiore et al., 2015). The interactions between air pollution and climate has been widely studied; however, there exist many uncertainties in the feedback loops (Fiore et al., 2015; IPCC, 2013; Tagaris et al., 2007; Weaver et al., 2009). In the United States, the EPA operates the IMPROVE network in remote areas, partially to capture background trends (Sisler and Malm, 2000). These, combined with satellite monitoring and regional and global air quality models, are important tools for characterizing changes in background concentrations.

Uncertainties exist in all observations and models. In accountability research, uncertainties inherent in typical emissions, air quality, exposure, and health outcome data and models increase with each assumption researchers make regarding causality. Uncertainties in epidemiological and other models come in two forms: stochastic, or sampling uncertainty, and uncertainty relating to model selection (Health Effects Institute, 2003). Scientists have formal methods for quantifying the former, but the latter is often

larger and more difficult to quantify. Overall uncertainty in estimates is an important aspect of interpreting results. Uncertainties can bias results and differential error can lead to erroneous conclusions when comparing effectiveness across regulations. More work is needed to better quantify uncertainties, and, when possible, correct for errors.

A final aspect of accountability studies is the comparison between the results of the study and the initial proposed benefits of the regulation (Figure 2-1, dashed lines). This is particularly important as policy makers use results of accountability studies to assess which types of regulatory programs are most effective.

2.8 Conclusions

As a research field, air pollution accountability has grown in complexity both in terms of the scales of regulations investigated and the analyses performed. The accountability chain serves as one framework for tracking the effects of regulatory actions across multiple outcomes, and future studies should continue to use it as a road map for assessing regulatory actions. An second framework, direct accountability, offers an alternative that should continue to be developed and complement traditional methods. Studies using either method need to account for relevant confounders when assigning baseline levels and estimating uncertainties with their results. New studies will be enhanced by planning appropriate studies before regulations are introduced in order to enhance data collection of confounding variables as well as outcomes. Further, new tools available to the scientific community, such as satellite measurements and statistical methods that have been applied in other domains offer broad opportunities for the expanded characterization of the efficacy of regulations. These should be evaluated with care in this domain, and applied as appropriate. The final step in an accountability should be to compare the results with the proposed benefits of the regulation of interest.

2.9 Acknowledgements

This material is based upon work supported by Health Effects Institute and the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1148903. Laura E. Carpenter helped design Figure 2-1.

2.10 References

Anderson, D.C., Loughner, C.P., Diskin, G., Weinheimer, A., Canty, T.P., Salawitch, R.J., Worden, H.M., Fried, A., Mikoviny, T., Wisthaler, A., Dickerson, R.R., 2014. Measured and modeled CO and NOy in DISCOVER-AQ: An evaluation of emissions and chemistry over the eastern US. Atmos. Environ. 96, 78–87. doi:10.1016/j.atmosenv.2014.07.004

Atkinson, R.W., Barratt, B., Armstrong, B., Anderson, H.R., Beevers, S.D., Mudway, I.S., Green, D., Derwent, R.G., Wilkinson, P., Tonne, C., 2009. The impact of the congestion charging scheme on ambient air pollution concentrations in London. Atmos. Environ. 43, 5493–5500. doi:10.1016/j.atmosenv.2009.07.023

Balachandran, S., Pachon, J.E., Hu, Y., Lee, D., Mulholland, J. a., Russell, A.G., 2012. Ensemble-trained source apportionment of fine particulate matter and method uncertainty analysis. Atmos. Environ. 61, 387–394. doi:10.1016/j.atmosenv.2012.07.031

Ban-Weiss, G.A., McLaughlin, J.P., Harley, R.A., Lunden, M.M., Kirchstetter, T.W., Kean, A.J., Strawa, A.W., Stevenson, E.D., Kendall, G.R., 2008. Long-term changes in emissions of nitrogen oxides and particulate matter from on-road gasoline and diesel vehicles. Atmos. Environ. 42, 220–232. doi:10.1016/j.atmosenv.2007.09.049

Bell, M.L., Morgenstern, R.D., Harrington, W., 2011. Quantifying the human health benefits of air pollution policies: Review of recent studies and new directions in accountability research. Environ. Sci. Policy 14, 357–368. doi:10.1016/j.envsci.2011.02.006

Blanchard, C.L., Tanenbaum, S., Hidy, G.M., 2012. Source contributions to atmospheric gases and particulate matter in the southeastern United States. Environ. Sci. Technol. 46, 5479–88. doi:10.1021/es203568t

Butler, T.J., Vermeylen, F.M., Rury, M., Likens, G.E., Lee, B., Bowker, G.E., McCluney, L., 2011. Response of ozone and nitrate to stationary source NOx emission reductions in the eastern USA. Atmos. Environ. 45, 1084–1094. doi:10.1016/j.atmosenv.2010.11.040

Camalier, L., Cox, W., Dolwick, P., 2007. The effects of meteorology on ozone in urban areas and their use in assessing ozone trends. Atmos. Environ. 41, 7127–7137. doi:10.1016/j.atmosenv.2007.04.061

Chang, H.H., Hao, H., Sarnat, S.E., 2014. A statistical modeling framework for projecting future ambient ozone and its health impact due to climate change. Atmos. Environ. 89, 290–297. doi:10.1016/j.atmosenv.2014.02.037

Chay, K., Greenstone, M., 2003. The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession. Q. J. Econ. 1121–1167.

Chen, Y., Ebenstein, A., Greenstone, M., Li, H., 2013. Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China ’s Huai River Policy. Proc. Natl. Acad. Sci. U. S. A. 110, 12936–12941. doi:10.1073/pnas.1300018110/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1300018110

Clancy, L., Goodman, P., Sinclair, H., Dockery, D.W., 2002. Effect of air-pollution control on death rates in Dublin, Ireland: an intervention study. Lancet 360, 1210–4. doi:10.1016/S0140-6736(02)11281-5

Colette, A., Granier, C., Hodnebrog, Ø., Jakobs, H., Maurizi, A., Nyiri, A., Bessagnet, B., D'Angiola, A., D'Isidoro, M., Gauss, M., Meleux, F., Memmesheimer, M., Mieville, A., Rouïl, L., Russo, F., Solberg, S., Stordal, F., Tampieri, F., 2011. Air quality trends in Europe over the past decade: a first multi-model assessment. Atmos. Chem. Phys. 11, 11657–11678. doi:10.5194/acp-11-11657-2011

Cox, L.A., 2013. Improving causal inferences in risk analysis. Risk Anal. 33, 1762–1771. doi:10.1111/risa.12072

Cox, L.A. (Tony), Popken, D. a., 2015. Has reducing fine particulate matter and ozone caused reduced mortality rates in the United States? Ann. Epidemiol. 25, 162–173. doi:10.1016/j.annepidem.2014.11.006

Cox, W.M., Chu, S.-H., 1993. METEOROLOGICALLY ADJUSTED OZONE TRENDS IN URBAN AREAS : A PROBABILISTIC APPROACH 27, 425–434.

Dallmann, T.R., Harley, R. a., 2010. Evaluation of mobile source emission trends in the United States. J. Geophys. Res. 115, D14305. doi:10.1029/2010JD013862

Daskalakis, N., Tsigaridis, K., Myriokefalitakis, S., Fanourgakis, G.S., Kanakidou, M., 2016. The success of emissions control legislation in mitigating air pollution is higher than previously estimated. Atmos. Chem. Phys. Discuss. 1–21. doi:10.5194/acp-2016-63

Dockery, D., Rich, D., George, P., Goodman, P., Ohman-Strickland, P., Clancy, L., Kotlov, T., 2013. Effect of Air Pollution Control on Mortality and Hospital Admissions in Ireland. doi:10.1097/01.ede.0000362914.17454.32

Dockery, D.W., Pope, C.A., Xu, X., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris, B.G., Speizer, F.E., 1993. An association between air pollution and mortality in six U.S. cities. N. Engl. J. Med. 329, 1753–1759.

Fiore, A.M., Naik, V., Leibensperger, E.M., Bachmann, J.D., 2015. Air Quality and Climate Connections. J. Air Waste Manage. Assoc. 65, 645–685. doi:10.1080/10962247.2015.1040526

Flaum, J.B., Rao, S.T., Zurbenko, I.G., 1996. Moderating the Influence of Meteorological Conditions on Ambient Ozone Concentrations. J. Air Waste Manage. Assoc. 46, 35–46. doi:10.1080/10473289.1996.10467439

Foley, K.M., Dolwick, P., Hogrefe, C., Simon, H., Timin, B., Possiel, N., 2015a. Dynamic evaluation of CMAQ part II: Evaluation of relative response factor metrics for ozone attainment demonstrations. Atmos. Environ. 103, 188–195. doi:10.1016/j.atmosenv.2014.12.039

Foley, K.M., Hogrefe, C., Pouliot, G., Possiel, N., Roselle, S.J., Simon, H., Timin, B., 2015b. Dynamic evaluation of CMAQ part I: Separating the effects of changing emissions and changing meteorology on ozone levels between 2002 and 2005 in the eastern US. Atmos. Environ. 103, 247–255. doi:10.1016/j.atmosenv.2014.12.038

Friedman, M., Powell, K., Hutwagner, L., Graham, L.M., Teague, W.G., 2001. Impact of changes in transportation and commuting behaviors during the 1996 Summer Olympic Games in Atlanta on air quality and childhood asthma. Jama 285, 897–905.

Garcia, V.C., Gego, E., Lin, S., Pantea, C., Rappazzo, K., Wootten, A., Rao, S.T., 2011. An evaluation of transported pollution and respiratory-related hospital admissions in the state of New York. Atmos. Pollut. Res. 2, 9–15. doi:10.5094/apr.2011.002

Gégo, E., Gilliland, A., Godowitch, J., Rao, S., Porter, P., Hogrefe, C., 2008. Modeling Analyses of the Effects of Changes in Nitrogen Oxides Emissions from the Electric Power Sector on Ozone Levels in the Eastern United States. J. Air Waste Manage. Assoc. 58, 580–588. doi:10.3155/1047-3289.58.4.580

Gégo, E., Porter, P.S., Gilliland, A., Rao, S.T., 2007. Observation-Based Assessment of the Impact of Nitrogen Oxides Emissions Reductions on Ozone Air Quality over the Eastern United States. J. Appl. Meteorol. Climatol. 46, 994–1008. doi:10.1175/JAM2523.1

Georgia EPD, 2013. Rules for Air Quality Control.

Godowitch, J.M., Gilliland, a. B., Draxler, R.R., Rao, S.T., 2008. Modeling assessment of point source NOx emission reductions on ozone air quality in the eastern United States. Atmos. Environ. 42, 87–100. doi:10.1016/j.atmosenv.2007.09.032

Godowitch, J.M., Pouliot, G. a., Trivikrama Rao, S., 2010. Assessing multi-year changes in modeled and observed urban NOX concentrations from a dynamic model evaluation perspective. Atmos. Environ. 44, 2894–2901. doi:10.1016/j.atmosenv.2010.04.040

Goldberg, D.L., Vinciguerra, T.P., Anderson, D.C., Hembeck, L., Canty, T.P., Ehrman, S.H., Martins, D.K., Stauffer, R.M., Thompson, A.M., Salawitch, R.J., Dickerson, R.R., 2016. CAMx Ozone Source Attribution in the Eastern United States using Guidance from Observations during DISCOVER-AQ Maryland. Geophys. Res. Lett. 2249–2258. doi:10.1002/2015GL067332

Goodman, P.G., Rich, D.Q., Zeka, a, Clancy, L., Dockery, D.W., 2009. Effect of Air Pollution Controls on Black Smoke and Sulfur Dioxide Concentrations across Ireland. J. Air Waste Manage. Assoc. 59, 207–213. doi:10.3155/1047-3289.59.1.207

Greenstone, M., 2004. Did the Clean Air Act cause the remarkable decline in sulfur dioxide concentrations? J. Environ. Econ. Manage. 47, 585–611. doi:10.1016/j.jeem.2003.12.001

Hao, N., Valks, P., Loyola, D., Cheng, Y.F., Zimmer, W., 2011. Space-based measurements of air quality during the World Expo 2010 in Shanghai. Environ. Res. Lett. 6, 044004. doi:10.1088/1748-9326/6/4/044004

Harley, R.A., 2014. On-road measurement of emissions from heavy-duty diesel trucks: impacts of fleet turnover and ARB's Drayage Truck Regulation.

Harrington, W., Morgenstern, R., Shih, J.-S., Bell, M.L., 2012. Did the Clean Air Act Amendments of 1990 really improve air quality? Air Qual. Atmos. Heal. 5, 353–367. doi:10.1007/s11869-012-0176-5

He, H., Stehr, J.W., Hains, J.C., Krask, D.J., Doddridge, B.G., Vinnikov, K.Y., Canty, T.P., Hosley, K.M., Salawitch, R.J., Worden, H.M., Dickerson, R.R., 2013. Trends in emissions and concentrations of air pollutants in the lower troposphere in the Baltimore/Washington airshed from 1997 to 2011. Atmos. Chem. Phys. 13, 7859–7874. doi:10.5194/acp-13-7859-2013

Health Effects Institute, 2003. Assessing Health Impact of Air Quality Regulations: Concepts and Methods for Accountability Research.

Hedley, A.J., Wong, C.-M., Thach, T.Q., Ma, S., Lam, T.-H., Anderson, H.R., 2002. Cardiorespiratory and all-cause mortality after restrictions on sulphur content of fuel in Hong Kong: an intervention study. Lancet 360, 1646–52.

Henneman, L.R., Chang, H.H., Lavoue, D., Mulholland, J.A., Russell, A.G., n.d. Accountability assessment of regulatory impacts on ozone and PM2.5 concentrations using statistical and deterministic pollutant sensitivities. Submitted.

Henneman, L.R.F., Holmes, H.A., Mulholland, J.A., Russell, A.G., 2015. Meteorological Detrending of Primary and Secondary Pollutant Concentrations: Method Application and Evaluation Using Long-Term (2000-2012) Data in Atlanta. Atmos. Environ. 119, 201–210. doi:10.1016/j.atmosenv.2015.08.007

Hopke, P.K., 2016. A Review of Receptor Modeling Methods for Source Apportionment. J. Air Waste Manage. Assoc. 2247, 10962247.2016.1140693. doi:10.1080/10962247.2016.1140693

Hou, Q., An, X.Q., Wang, Y., Guo, J.P., 2010. An evaluation of resident exposure to respirable particulate matter and health economic loss in Beijing during Beijing 2008 Olympic Games. Sci. Total Environ. 408, 4026–32. doi:10.1016/j.scitotenv.2009.12.030

Hu, Y., Balachandran, S., Pachon, J.E., Baek, J., Ivey, C., Holmes, H., Odman, M.T., Mulholland, J. a., Russell, a. G., 2014. Fine particulate matter source apportionment using a hybrid chemical transport and receptor model approach. Atmos. Chem. Phys. 14, 5415–5431. doi:10.5194/acp-14-5415-2014

Huang, K., Zhuang, G., Lin, Y., Wang, Q., Fu, J.S., Fu, Q., Liu, T., Deng, C., 2013. How to improve the air quality over megacities in China: Pollution characterization and source analysis in Shanghai before, during, and after the 2010 World Expo. Atmos. Chem. Phys. 13, 5927–5942. doi:10.5194/acp-13-5927-2013

Huang, X.F., He, L.Y., Xue, L., Sun, T.L., Zeng, L.W., Gong, Z.H., Hu, M., Zhu, T., 2012. Highly time-resolved chemical characterization of atmospheric fine particles during 2010 Shanghai World Expo. Atmos. Chem. Phys. 12, 4897–4907. doi:10.5194/acp-12-4897-2012

Hubbell, B., Greenbaum, D., 2014. Counterpoint: Moving From Potential-Outcomes Thinking to Doing--Changing Research Planning to Enable Successful Health Outcomes Research. Am. J. Epidemiol. 180, 1141–1144. doi:10.1093/aje/kwu266

Hubbell, B.J., Crume, R. V., Evarts, D.M., Cohen, J.M., 2009. Policy Monitor: Regulation and Progress under the 1990 Clean Air Act Amendments. Rev. Environ. Econ. Policy 4, 122–138. doi:10.1093/reep/rep019

IPCC 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, 2013. doi:doi:10.1017/CBO9781107415324

Kang, D., Hogrefe, C., Foley, K.L., Napelenok, S.L., Mathur, R., Trivikrama Rao, S., 2013. Application of the Kolmogorov–Zurbenko filter and the decoupled direct 3D method for the dynamic evaluation of a regional air quality model. Atmos. Environ. 80, 58–69. doi:10.1016/j.atmosenv.2013.04.046

Kean, A.J., Littlejohn, D., Ban-Weiss, G.A., Harley, R.A., Kirchstetter, T.W., Lunden, M.M., 2009. Trends in on-road vehicle emissions of ammonia. Atmos. Environ. 43, 1565–1570. doi:10.1016/j.atmosenv.2008.09.085

Kelly, F., Anderson, H.R., Armstrong, B., Atkinson, R.W., Barratt, B., Beevers, S., Derwent, D., Green, D., Mudway, I., Wilkinson, P., 2011. The Impact of the Congestion Charging Scheme on Air Quality in London.

Klausbruckner, C., Annegarn, H.J., Henneman, L.R.F., Rafaj, P., 2016. A policy review of synergies and trade-offs in South African climate change mitigation and air pollution control strategies. Environ. Sci. Policy 57, 70–78. doi:10.1016/j.envsci.2015.12.001

Kuebler, J., van den Bergh, H., Russell, A.G., 2001. Long-term trends of primary and secondary pollutant concentrations in Switzerland and their response to emission controls and economic changes. Atmos. Environ. 35, 1351–1363. doi:10.1016/S1352-2310(00)00401-5

Kuwayama, T., Schwartz, J.R., Harley, R. a., Kleeman, M.J., 2013. Particulate Matter Emissions Reductions due to Adoption of Clean Diesel Technology at a Major Shipping Port. Aerosol Sci. Technol. 47, 29–36. doi:10.1080/02786826.2012.720049

Laden, F., Schwartz, J., Speizer, F.E., Dockery, D.W., 2006. Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study. Am. J. Respir. Crit. Care Med. 173, 667–72. doi:10.1164/rccm.200503-443OC

Lamsal, L.N., Martin, R. V., Padmanabhan, A., Van Donkelaar, A., Zhang, Q., Sioris, C.E., Chance, K., Kurosu, T.P., Newchurch, M.J., 2011. Application of satellite observations for timely updates to global anthropogenic NOx emission inventories. Geophys. Res. Lett. 38, 1–5. doi:10.1029/2010GL046476

Lawson, D., Groblicki, P., Stedman, D., Bishop, G., Guenther, P., 1990. Emissions of In-Use motor vehicle in Los Angeles: A pilot study of remote sensing and the inspection and maintenance program. J. Air Waste Manag. Assoc. 40, 1096 – 1105. doi:10.1080/10473289.1990.10466754

Lawson, D.R., 1990. The Southern California Air Quality Study. J. Air Waste Manage. Assoc. 40, 156–165. doi:10.1080/10473289.1990.10466671

Lee, J.-T., Son, J.-Y., Cho, Y.-S., 2007. Benefits of Mitigated Ambient Air Quality Due to Transportation Control on Childhood Asthma Hospitalization during the 2002 Summer Asian Games in Busan, Korea. J. Air Waste Manage. Assoc. 57, 968–973. doi:10.3155/1047-3289.57.8.968

Lee, M., Kloog, I., Chudnovsky, A., Lyapustin, A., Wang, Y., Melly, S., Coull, B., Koutrakis, P., Schwartz, J., 2015. Spatiotemporal prediction of fine particulate matter using high-resolution satellite images in the Southeastern US 2003-2011. J. Expo. Sci. Environ. Epidemiol. doi:10.1038/jes.2015.41

Li, Y., Wang, W., Kan, H., Xu, X., Chen, B., 2010. Air quality and outpatient visits for asthma in adults during the 2008 Summer Olympic Games in Beijing. Sci. Total Environ. 408, 1226–7. doi:10.1016/j.scitotenv.2009.11.035

Lim, S.S., Vos, T., Flaxman, A.D., et al., 2013. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380, 2224–60. doi:10.1016/S0140-6736(12)61766-8

Lin, J., Nielsen, C.P., Zhao, Y., Lei, Y., Liu, Y., McElroy, M.B., 2010. Recent changes in particulate air pollution over china observed from space and the ground: Effectiveness of emission control. Environ. Sci. Technol. 44, 7771–7776. doi:10.1021/es101094t

Lin, S., Jones, R., Pantea, C., Özkaynak, H., Rao, S.T., Hwang, S.-A., Garcia, V.C., 2013. Impact of NOx emissions reduction policy on hospitalizations for respiratory disease in New York State. J. Expo. Sci. Environ. Epidemiol. doi:10.1038/jes.2012.69

Lin, Y., Huang, K., Zhuang, G., Fu, J.S., Xu, C., Shen, J., Chen, S., 2013. Air quality over the Yangtze River Delta during the 2010 Shanghai expo. Aerosol Air Qual. Res. 13, 1655–1666. doi:10.4209/aaqr.2012.11.0312

Liu, Y., He, K., Li, S., Wang, Z., Christiani, D.C., Koutrakis, P., 2012. A statistical model to evaluate the effectiveness of PM2.5 emissions control during the Beijing 2008 Olympic Games. Environ. Int. 44, 100–5. doi:10.1016/j.envint.2012.02.003

Liu, Y., Yuan, B., Li, X., Shao, M., Lu, S., Li, Y., Chang, C.C., Wang, Z., Hu, W., Huang, X., He, L., Zeng, L., Hu, M., Zhu, T., 2015. Impact of pollution controls in Beijing on atmospheric oxygenated volatile organic compounds (OVOCs) during the 2008 olympic games: Observation and modeling implications. Atmos. Chem. Phys. 15, 3045–3062. doi:10.5194/acp-15-3045-2015

Lurmann, F., Avol, E., Gilliland, F., 2014. Emissions reduction policies and recent trends in Southern California's ambient air quality. J. Air Waste Manage. Assoc. 65, 324–335. doi:10.1080/10962247.2014.991856

McDonald, B.C., Dallmann, T.R., Martin, E.W., Harley, R. a., 2012. Long-term trends in nitrogen oxide emissions from motor vehicles at national, state, and air basin scales. J. Geophys. Res. Atmos. 117, n/a–n/a. doi:10.1029/2012JD018304

McDonald, B.C., Gentner, D.R., Goldstein, A.H., Harley, R. a, 2013. Long-Term Trends in Motor Vehicle Emissions in U.S. Urban Areas. Environ. Sci. Technol. 47, 10022–31. doi:10.1021/es401034z

Milanchus, M.L., Trivikrama Rao, S., Zurbenko, I.G., 1998. Journal of the Air & Waste Management Association Evaluating the Effectiveness of Ozone Management Efforts in the Presence of Meteorological Variability Evaluating the Effectiveness of Ozone Management Efforts in the Presence of Meteorological Variability. J. Air Waste Manag. Assoc. J. Air Waste Manag. Assoc. J. Air Waste Manag. Assoc 483, 201–215. doi:10.1080/10473289.1998.10463673

Morgenstern, R.D., Harrington, W., Shih, J.-S., Bell, M.L., 2012. Accountability Analysis of Title IV Phase 2 of the 1990 Clean Air Act Amendments.

Murray, F., 2013. The changing winds of atmospheric environment policy. Environ. Sci. Policy 29, 115–123. doi:10.1016/j.envsci.2013.02.005

Napelenok, S.L., Foley, K.M., Kang, D., Mathur, R., Pierce, T., Rao, S.T., 2011. Dynamic evaluation of regional air quality model's response to emission reductions in the presence of uncertain emission inventories. Atmos. Environ. 45, 4091–4098. doi:10.1016/j.atmosenv.2011.03.030

National Research Council, 2004. Air Quality Management in the United States.

Pachon, J., Balachandran, S., 2012. Development of outcome-based, multipollutant mobile source indicators. J. Air … 11.

Peel, J.L., Klein, M., Flanders, W.D., Mulholland, J.A., Tolbert, P.E., 2010. Impact of Improved Air Quality During the 1996 Summer Olympic Games in Atlanta on Multiple Cardiovascular and Respiratory Outcomes.

Pegues, A.H., Cohan, D.S., Digar, A., Douglass, C., Wilson, R.S., 2012. Efficacy of recent state implementation plans for 8-hour ozone. J. Air Waste Manage. Assoc. 62, 252–261. doi:10.1080/10473289.2011.646049

Peters, A., Breitner, S., Cyrys, J., Stölzel, M., Pitz, M., Wölke, G., Heinrich, J., Kreyling, W., Küchenhoff, H., Wichmann, H.-E., 2009. The influence of improved air quality on mortality risks in Erfurt, Germany. Res. Rep. Health. Eff. Inst. 5–77; discussion 79–90. doi:1041-5505 (Print)

Peters, J., Hedley, A.J., Wong, C.M., Lam, T.H., Ong, S.G., Liu, J., Spiegelhalter, D.J., 1996. Effects of ambient air pollution and environmental tobacco smoke on respiratory health of non-smoking women in Hong Kong. Int. J. Epidemiol. 28, 859. doi:10.1093/ije/28.5.859

Pierce, T., Hogrefe, C., Trivikrama Rao, S., Porter, P.S., Ku, J.-Y., 2010. Dynamic evaluation of a regional air quality model: Assessing the emissions-induced weekly ozone cycle. Atmos. Environ. 44, 3583–3596. doi:10.1016/j.atmosenv.2010.05.046

Pierson, W.R., Gertler, A.W., Bradow, R.L., 1990. Comparison of the SCAQS Tunnel Study With Other On-Road Vehicle Emissions Data. J. Air Waste Manag. 40, 1495–1504. doi:10.1080/10473289.1990.10466799

Pope, C., 1989. Respiratory disease associated with community air pollution and a steel mill, Utah Valley. Am. J. Public Health 79, 623–8.

Pope, C.A., Cropper, M., Coggins, J., Cohen, A., 2015. Health Benefits of Air Pollution Abatement Policy: Role of the Shape of the Concentration-Response Function. J. Air Waste Manage. Assoc. 2247, 141217131236009. doi:10.1080/10962247.2014.993004

Pope, C.A., Ezzati, M., Dockery, D.W., 2009. Fine-particulate air pollution and life expectancy in the United States. N. Engl. J. Med. 360, 376–86. doi:10.1056/NEJMsa0805646

Pope, C.A., Ezzati, M., Dockery, D.W., 2012. Validity of observational studies in accountability analyses: The case of air pollution and life expectancy. Air Qual. Atmos. Heal. 5, 231–235. doi:10.1007/s11869-010-0130-3

Rao, S., Zalewsky, E., Zurbenko, I., 1995. Determining temporal and spatial variations in ozone air quality. J. Air Waste … 45, 57–61.

Rao, S., Zurbenko, I., 1994. Detecting and tracking changes in ozone air quality. Air waste 44, 1089–1092.

Rich, D.Q., Kipen, H.M., Huang, W., Wang, G., Wang, Y., Zhu, P., Ohman-Strickland, P., Hu, M., Philipp, C., Diehl, S.R., Lu, S.-E., Tong, J., Gong, J., Thomas, D., Zhu, T., Zhang, J. (Jim), 2012. Association between changes in air pollution levels during the Beijing Olympics and biomarkers of inflammation and thrombosis in healthy young adult. J. Am. Med. Assoc. 307, 2068–2078. doi:10.1001/jama.2012.3488

Rich, D.Q., Liu, K., Zhang, J., Thurston, S.W., Stevens, T.P., Pan, Y., Kane, C., Weinberger, B., Ohman-Strickland, P., Woodruff, T.J., Duan, X., Assibey-Mensah, V., Zhang, J., 2015. Differences in Birth Weight Associated with the 2008 Beijing Olympic Air Pollution Reduction: Results from a Natural Experiment. Environ. Health Perspect. 117, 1713–1717. doi:10.1289/ehp.1408795

Russell, M.C., 2015. The Impact of Three Recent Coal-fired Power Plant Closings on Pittsburgh Air Quality: A Natural Experiment. J. Air Waste Manage. Assoc. 2247. doi:10.1080/10962247.2016.1170738

Schreifels, J.J., Fu, Y., Wilson, E.J., 2012. Sulfur dioxide control in China: policy evolution during the 10th and 11th Five-year Plans and lessons for the future. Energy Policy 48, 779–789. doi:10.1016/j.enpol.2012.06.015

Sickles II, J.E., Shadwick, D.S., 2014. Air quality and atmospheric deposition in the eastern US: 20 years of change. Atmos. Chem. Phys. 14, 173–197. doi:10.5194/acp-15-173-2015

Simon, H., Baker, K.R., Akhtar, F., Napelenok, S.L., Possiel, N., Wells, B., Timin, B., 2013. A direct sensitivity approach to predict hourly ozone resulting from compliance with the national ambient air quality standard. Environ. Sci. Technol. 47, 2304–2313. doi:10.1021/es303674e

Simon, H., Baker, K.R., Phillips, S., 2012. Compilation and interpretation of photochemical model performance statistics published between 2006 and 2012. Atmos. Environ. 61, 124–139. doi:10.1016/j.atmosenv.2012.07.012

Simon, H., Reff, A., Wells, B., Xing, J., Frank, N., 2015. Ozone trends across the United States over a period of decreasing NOx and VOC emissions. Environ. Sci. Technol. 49, 186–95. doi:10.1021/es504514z

Sisler, J.F., Malm, W.C., 2000. Interpretation of Trends of PM2.5 and Reconstructed Visibility from the IMPROVE Network. J. Air Waste Manage. Assoc. 50, 775–789. doi:10.1080/10473289.2000.10464127

Souri, A.H., Choi, Y., Jeon, W., Li, X., Pan, S., Diao, L., Westenbarger, D.A., 2016. Constraining NOx emissions using satellite NO2 measurements during 2013 DISCOVER-AQ Texas campaign. Atmos. Environ. 131, 371–381. doi:10.1016/j.atmosenv.2016.02.020

Streets, D.G., Canty, T., Carmichael, G.R., De Foy, B., Dickerson, R.R., Duncan, B.N., Edwards, D.P., Haynes, J.A., Henze, D.K., Houyoux, M.R., Jacob, D.J., Krotkov, N.A., Lamsal, L.N., Liu, Y., Lu, Z., Martin, R. V., Pfister, G.G., Pinder, R.W., Salawitch, R.J., Wecht, K.J., 2013. Emissions estimation from satellite retrievals: A review of current capability. Atmos. Environ. 77, 1011–1042. doi:10.1016/j.atmosenv.2013.05.051 Review

Su, C., Hampel, R., Franck, U., Wiedensohler, A., Cyrys, J., Pan, X., Wichmann, H.E., Peters, A., Schneider, A., Breitner, S., 2015. Assessing responses of cardiovascular mortality to particulate matter air pollution for pre-, during- and post-2008 Olympics periods. Environ. Res. 142, 112–122. doi:10.1016/j.envres.2015.06.025

Tagaris, E., Manomaiphiboon, K., Liao, K.J., Leung, L.R., Woo, J.H., He, S., Amar, P., Russell, A.G., 2007. Impacts of global climate change and emissions on regional ozone and fine particulate matter concentrations over the United States. J. Geophys. Res. Atmos. 112. doi:10.1029/2006JD008262

U.S. EPA, 1997a. Regulatory Impact Analysis : Control of Emissions of Air Pollution from Highway Heavy-Duty Engines.

U.S. EPA, 1997b. The Benefits and Costs of the Clean Air Act, 1970 to 1990.

U.S. EPA, 1999a. The Benefits and Costs of the Clean Air Act 1990 to 2010 EPA Report to Congress.

U.S. EPA, 1999b. Regulatory Impact Analysis–Control of Air Pollution from New Motor Vehicles: Tier 2 Motor Vehicle Emissions Standards and Gasoline Sulfur Control Requirements, US Environmental Protection Agency, Air and ….

U.S. EPA, 2005a. Regulatory Impact Analysis of the Final Clean Air Mercury Rule.

U.S. EPA, 2005b. Regulatory Impact Analysis for the Final Clean Air Interstate Rule.

U.S. EPA, 2011. The Benefits and Costs of the Clean Air Act from 1990 to 2020 Final Report.

U.S. EPA, 2016. Air Markets Program Data [WWW Document]. URL https://ampd.epa.gov/ampd/

van der A, R.J., Mijling, B., Ding, J., Koukouli, M.E., Liu, F., Li, Q., Mao, H., Theys, N., 2016. Cleaning up the air: Effectiveness of air quality policy for SO2 and NOx emissions in China. Atmos. Chem. Phys. Discuss. 1–18. doi:10.5194/acp-2016-445

van Erp, A.M., Kelly, F.J., Demerjian, K.L., Pope, C.A., Cohen, A.J., 2011. Progress in research to assess the effectiveness of air quality interventions towards improving public health. Air Qual. Atmos. Heal. 5, 217–230. doi:10.1007/s11869-010-0127-y

Wang, B., Shao, M., Lu, S.H., Yuan, B., Zhao, Y., Wang, M., Zhang, S.Q., Wu, D., 2010. Variation of ambient non-methane hydrocarbons in Beijing city in summer 2008. Atmos. Chem. Phys. 10, 5911–5923. doi:10.5194/acp-10-5911-2010

Wang, M., Zhu, T., Zheng, J., Zhang, R.Y., Zhang, S.Q., Xie, X.X., Han, Y.Q., Li, Y., 2009. Use of a mobile laboratory to evaluate changes in on-road air pollutants during the Beijing 2008 Summer Olympics. Atmos. Chem. Phys. 9, 12857–12898. doi:10.5194/acpd-9-12857-2009

Wang, S., Xing, J., Zhao, B., Jang, C., Hao, J., 2014. Effectiveness of national air pollution control policies on the air quality in metropolitan areas of China. J. Environ. Sci. 26, 13–22. doi:10.1016/S1001-0742(13)60381-2

Wang, S., Zhao, M., Xing, J., Wu, Y., Zhou, Y., Lei, Y., He, K., Fu, L., Hao, J., 2010. Quantifying the air pollutants emission reduction during the 2008 olympic games in Beijing. Environ. Sci. Technol. 44, 2490–2496. doi:10.1021/es9028167

Wang, T., Xie, S., 2009. Assessment of traffic-related air pollution in the urban streets before and during the 2008 Beijing Olympic Games traffic control period. Atmos. Environ. 43, 5682–5690. doi:10.1016/j.atmosenv.2009.07.034

Wang, W., Primbs, T., Tao, S., Simonich, S.L.M., 2009. Atmospheric Particulate Matter Pollution during the 2008 Beijing Olympics. Environ. Sci. Technol. 43, 5314–5320. doi:10.1029/2006JD007448.Xiaohong

Wang, X., Westerdahl, D., Chen, L.C., Wu, Y., Hao, J., Pan, X., Guo, X., Zhang, K.M., 2009. Evaluating the air quality impacts of the 2008 Beijing Olympic Games: On-road emission factors and black carbon profiles. Atmos. Environ. 43, 4535–4543. doi:10.1016/j.atmosenv.2009.06.054

Wang, Y., Hao, J., McElroy, M.B., Munger, J.W., Ma, H., Chen, D., Nielsen, C.P., 2009. Ozone air quality during the 2008 Beijing Olympics – effectiveness of emission restrictions. Atmos. Chem. Phys. Discuss. 9, 9927–9959. doi:10.5194/acpd-9-9927-2009

Weaver, C.P., Liang, X.Z., Zhu, J., et al., 2009. A preliminary synthesis of modeled climate change impacts on U.S. regional ozone concentrations. Bull. Am. Meteorol. Soc. 90, 1843–1863. doi:10.1175/2009BAMS2568.1

Wong, C.-M., Rabl, A., Thach, T.Q., Chau, Y.K., Chan, K.P., Cowling, B.J., Lai, H.K., Lam, T.H., McGhee, S.M., Anderson, H.R., Hedley, A.J., 2012. Impact of the 1990 Hong Kong legislation for restriction on sulfur content in fuel.

Wong, C.M., Lam, T.H., Peters, J., Hedley, A.J., Ong, S.G., Tam, A.Y., Liu, J., Spiegelhalter, D.J., 1998. Comparison between two districts of the effects of an air pollution intervention on bronchial responsiveness in primary school children in Hong Kong. J. Epidemiol. Community Health 52, 571–578. doi:10.1136/jech.52.9.571

Wu, S., Deng, F., Niu, J., Huang, Q., Liu, Y., Guo, X., 2010. Association of heart rate variability in taxi drivers with marked changes in particulate air pollution in Beijing in 2008. Environ. Health Perspect. 118, 87–91. doi:10.1289/ehp.0900818

Wyat Appel, K., Bhave, P. V., Gilliland, A.B., Sarwar, G., Roselle, S.J., 2008. Evaluation of the community multiscale air quality (CMAQ) model version 4.5: Sensitivities impacting model performance; Part II—particulate matter. Atmos. Environ. 42, 6057–6066. doi:10.1016/j.atmosenv.2008.03.036

Xing, J., Zhang, Y., Wang, S., Liu, X., Cheng, S., Zhang, Q., Chen, Y., Streets, D.G., Jang, C., Hao, J., Wang, W., 2011. Modeling study on the air quality impacts from emission reductions and atypical meteorological conditions during the 2008 Beijing Olympics. Atmos. Environ. 45, 1786–1798. doi:10.1016/j.atmosenv.2011.01.025

Yuan, Z., Yadav, V., Turner, J.R., Louie, P.K.K., Lau, A.K.H., 2013. Long-term trends of ambient particulate matter emission source contributions and the accountability of control strategies in Hong Kong over 1998-2008. Atmos. Environ. 76, 21–31. doi:10.1016/j.atmosenv.2012.09.026

Zigler, C., Dominici, F., 2016. Demystifying Causal Inference In Air Pollution Epidemiology, in: Health Effects Institute Annual Conference. Denver, CO.

Zigler, C.M., Dominici, F., 2014. Point: Clarifying Policy Evidence With Potential-Outcomes Thinking--Beyond Exposure-Response Estimation in Air Pollution Epidemiology. Am. J. Epidemiol. 180, 1133–1140. doi:10.1093/aje/kwu263

Zigler, C.M., Dominici, F., Wang, Y., 2012. Estimating causal effects of air quality regulations using principal stratification for spatially correlated multivariate intermediate outcomes. Biostatistics 13, 289–302. doi:10.1093/biostatistics/kxr052

Zigler, C.M., Kim, C., Choirat, C., Hansen, J.B., Wang, Y., Hund, L., Samet, J., Dominici, F., 2016. Causal Inference Methods for Estimating Long-Term Health Effects of Air Quality Regulations.

CHAPTER 3. REGULATORY IMPACTS ON POWER PLANT AND MOBILE SOURCE EMISSIONS IN THE SOUTHEASTERN UNITED STATES[2]

3.1 Abstract

Electric utilities and automobiles emit vast amounts of air pollutants. These sources, however, are heavily regulated, and their emissions have been curbed, particularly since the 1970 Clean Air Act Amendments in the United States. Previous studies have linked air pollution regulations to large changes over time; however, if the goal is to dissect emission changes and attribute reductions to *specific* regulatory actions to assess which previous regulations have been more effective than others and thus plan future regulations, a broad-brushed before-after comparison will not suffice. Other factors, such as fuel price, population shifts, are all linked, and should be considered in this type of assessment. This work inspects electricity generating unit emissions from the Environmental Protection Agency's (EPA) Air Markets Program Database and mobile source emissions modelled using EPA's Motor vehicle emissions simulator (MOVES) model in Atlanta, GA and the broader Southeast. The approach assesses emissions and demand changes, and links emissions changes to multiple regulations from each source type after 1990. The output is daily counterfactual (i.e., hypothetical scenarios assuming various regulatory programs were not implemented) time series for multiple regulatory programs. Counterfactuals show increasing effectiveness across years as programs are incrementally implemented, and varying impacts of different programs on various pollutants.

3.2 Introduction

The 1990 Clean Air Act Amendments (CAAA) marked a turning point in regulating air quality across the United States. The 1970 CAAA (which amended the original 1963 Act) gave the United States Environmental Protection Agency (EPA) the authority to regulate air pollutants using two specific tools: air quality standards and emissions limits

[2] Portions of this chapter are due for publication in: Russell, Armistead G, Paige E Tolbert, Lucas RF Henneman, Joe Abrams, Cong Liu, Mitch Klein, James A Mulholland, et al. 2017. "Impacts of Regulations on Air Quality and Emergency Department visits in the Atlanta metropolitan area, 1999-2013" *In press*.

(National Research Council, 2004). EPA sets National Ambient Air Quality Standards (NAAQS) for six "criteria" pollutants—ozone (O_3), particulate matter (PM), carbon monoxide (CO), sulfur dioxide (SO_2), nitrogen dioxide (NO_2), and lead—at levels designed to protect public human health and public welfare. Each pollutant has both primary (health) and secondary (welfare) standards (the same for many species), and PM is regulated both as $PM_{2.5}$ (particle diameters less than 2.5 μm) and as PM_{10} (diameters less than 10 μm). NAAQS are written as concentrations averaged over a specific period and follow specific statistical forms unique to each pollutant. EPA designates areas in exceedance of the NAAQS as non-attainment areas (NAAs), and requires the encompassing state to develop a State Implementation Plan (SIP) for reducing ambient air quality concentrations below the standards. The 1990 CAAA clarified and expanded the EPA's previous authority related to NAAQS-setting and enforcement, mobile and stationary source emissions standards, emissions cap-and-trade programs, and permit requirements. The Amendments also established the EPA's jurisdiction to regulate air toxics (hazardous air pollutants) and chemicals related to the stratospheric O_3 depletion.

Emission standards aim to reduce the release of air pollutants from specific industries and source types, and are written either as emission rates (emission per activity, e.g., grams NO_X mi^{-1}) or as total allowable emissions over a specified amount of time. Some standards are applied to specific plants, while others are applied to a fleet, and some regulatory programs—e.g., the Acid Rain Program defined in the 1990 CAAA—set up trading markets that permit plant owners to buy and sell emission allowances (National Research Council, 2004). For mobile sources, recent regulatory programs—e.g., the Tier 2 Gasoline Vehicle Standards and the 2007 Heavy Duty Diesel Rule—set standards for both engines and fuel composition. Mobile-source emissions limits are set at a federal level; however, EPA allows one state—California—to set mobile emissions standards independent of the national levels (though they must be at least as stringent), and other states can adopt either the federal or California standards. States use other tools, such as limits on Reid Vapor Pressure in gasoline below federal limits and Inspection and Maintenance (IM) programs, to reduce emissions in NAAs.

In response to regulations contained in the 1990 CAAA, the EPA and the Georgia Department of Natural Resources' Environmental Protection Division (EPD) have applied

various regulatory tools to improve air quality, with a focus on Atlanta, which frequently exceeds the NAAQS for O_3 and $PM_{2.5}$. Assessments of the effectiveness of specific regulations, however, are made difficult by the complex interplay between national regulations and their implementation at the state and local levels. For example, the EPD has implemented several emissions standards on stationary sources separate from national programs. Often, the state programs—codified in SIPs—are similar in approach and timing to national programs, and may be developed in negotiations between regulators, utilities, and public service commissions (PSC) that govern utilities. A state may promulgate a rule to achieve multiple objectives or meet multiple national standards (e.g., O_3 and $PM_{2.5}$ share precursors). Further, a utility, whose actions are subject to PSC rulings, may seek to identify the most cost effective measures to address multiple regulations. Such interconnected emissions policies affect air quality in varying ways depending on multiple factors such as source industry, location, and stack height, economic activity, and climate. Any assessment of the effectiveness of specific regulations implemented under the CAAA, therefore, must begin by acknowledging the intermeshed nature of air pollution regulations.

3.3 Power plant emissions

This section describes the emissions reductions of electricity generating unit (EGU) nitrogen oxides ($NO_X = NO + NO_2$) and sulphur dioxide (SO_2) over the 20-year period from 1995-2014 in Atlanta. EGU emissions are regulated under several policies, both national and state. It is of interest to regulators and stakeholders to assess the response of emissions to specific regulatory actions. This information can be used to link air quality and health improvements to regulations in and accountability framework (Health Effects Institute, 2003). This report investigates 20 years of emissions data and regulatory actions, and seeks to attribute emissions reductions to specific policies enacted over this period.

3.3.1 Decision-making by electric utilities under the current regulatory framework

Electric utilities are in a unique position. Many are public companies, meaning they are obligated to maximize shareholder value. However, their profit is limited by the amount that they can charge for electricity, which is regulated by the local Public Services Commission (PSC). Each investor-owned utility (Georgia Power is the only such utility in Georgia) submits an integrated resource plan (IRP) every three years to the PSC for approval. The IRP is the company's 20-year outlook, and includes information on

electricity supply/demand projections, the regulatory outlook, control options, and options for commissioning/shuttering new/old plants.

There are four major drivers of these actions by major electric (Ewald, 2015):

Compliance. Units must comply with existing regulations to operate. However, utilities will work to ensure new regulatory actions are achievable.

Cost to the consumer. Utilities work with the PSC to create a pricing structure that is in the consumer's best interest.

Avoid new source review. The national new source review (NSR) program is a process that assesses the emissions from new sources. Repermitting occurs every 5-8 years, and can cover existing plants that show a 'significant modification,' which can include running the plant more. One way to avoid NSR is to show that a modification to a plant will result in reduced emissions.

Co-owners. Many plants are owned by multiple utilities (this is the case of multiple plants in Georgia). The other utilities are either electric membership corporations (EMCs) or municipalities, neither of which have their prices set by the PSC. The differences between the business models of the three types of electricity-producers creates difficulties in negotiating who will pay for controls, plant lifetimes, etc.

Profit. Utilities must make a profit to stay solvent. The PSC considers necessary controls (both the installation and the cost of running the control once it is installed), new plants, maintenance, etc. when setting the price.

Cost recovery is not an option when controls are not required under existing or anticipated legislation (Ewald, 2015; Huling, 2014). For this research, this means that controls on plants should all be able to be linked to individual actions by the state or federal governments. This is a key assumption in the following assessment.

Utilities face multiple time-scales of decisions making in terms of how to provide electricity to their clients. On the shortest scale, the grid needs to stay balanced (supply must equal demand). At this scale, decisions include when scale up/down certain plants to keep the energy supply consistent with the demand. At longer time scales, decisions are made regarding fuel switching, building new capacity, retiring existing plants, and adding controls.

Because of this, it is difficult to bound a group of plants into divisions based on the demand they serve. Utilities will produce, buy, and export power based on the total demand of the locations they serve, the cost of producing power in their own plants, and the cost of purchasing power from other utilities. Units across the fleet are `stacked'—in terms of their status as base, mid, or peaking load—in a way that minimizes the cost per kW-hr produced. With these in mind, utilities assign load across their fleet of plants. Depending on the price of different fuels and the availability of plants, load may be covered by plants that are nearby, in a different part of the state, or in a nearby state. In general, plants that are heavily controlled tend to run more.

As emission regulations have been put into place, the flexibility of being able to switch between plants has decreased (Ewald, 2015; Huling, 2014). Utilities create averaging plans that set a maximum emissions factor for a fleet of plans, allows for some flexibility for plant management. In non-attainment areas (NAAs), states designate specific plants for emissions reductions to meet NAAQS standards. Utilities may be forced to rely more or less heavily on these plants in order to satisfy fleet averaging plans or reduce emissions to standards set in the State Implementation Plan.

3.3.2 Power plants included in the study

To study the effects of regulatory actions that effect emissions from EGUs on in the Atlanta, GA region, we selected only the major (greater than 25 mega-watt equivalents (MWe)) plants within the 20-county Atlanta $PM_{2.5}$ non-attainment are (ANAA) that were in operation for all or part of the time span from 1995 - 2013. EGUs used in the study are shown in Figure 5-1. Reasons for selecting these plants (Atkinson, Bowen, Chattahoochee, Doyle, Harlley Branch, Hawk Road, McDonough, MPC, Tenaska, Walton, Wansley & Yates) are:

- Daily plant-level emissions are publically available in CEM data from the EPA
- Plants within the ANAA have been regulated most strictly over the study period compared with others in Georgia
- Emissions from these plants assumed to have a relatively larger impact on air quality on the central monitor in Atlanta (used for health analyses) than other facilities in the region (Muller, Tong, & Mendelsohn, 2009)

- Long-term emissions reductions reflect similar changes observed across the southeastern US (U.S. EPA, 2016)

CEM data, downloaded from the EPA Air Markets Database, includes notations on when certain controls went into operation on different plants. This information is synthesized for each of the plants that operated over the study period in the NAA (Table 3-1). Some of the plants were not online for the entirety of the study period, and some information in this section (including each plant's status as a base load or a peaking plant) is inferred from the record of load in the CEM data.

Table 3-1. Controls on plants used in the analysis.

Plant	Year Built	Primary Fuel	NO_x				SO_2		$PM_{2.5}$		Fuel Retrofit	Retrofit Date	Notes
			Control 1	Install Date	Control 2	Install Date	Control 1	Install Date	Control 1	Install Date			
Atkinson	1930	Coal											
Bowen	1975	Coal	Low NO_x burners	1995	SCR[1]	2001 (2 units) 2003 (2 units)	FGD[2]	2008-2010	Electrostatic precipitator	pre-1990			
Chattahoochee	2003	Natural Gas	SCR[1]	2003									
Doyle	2000	Natural Gas	Dry Low NO_x burners	2000									
Harllee Branch	1965	Coal	Low NO_x burners	2002 (2 units) 2003 (2 units)					Electrostatic precipitator	pre-1990			
Hawk Road	2001	Natural Gas	Dry Low NO_x burners	2001									
Jack McDonough	1963	Coal	Low NO_x burners w/ separated OFA	1995	Natural Gas Co-firing	1999			Electrostatic precipitator	pre-1990	Natural Gas	2011 (2 units) 2012 (4 units)	Retrofit included Dry Low NO_x burners and SCRs[1]
MPC	1999	Natural Gas	Water Injection	1999									
Tenaska	2001	Natural Gas	Dry Low NO_x burners	2001	Water Injection	2001							
Walton	2001	Natural Gas	Dry Low NO_x burners	2001									
Wansley (6052)	1976	Coal	Low NO_x burners w/ Closed-coupled/ separated OFA	1995	SCR[1]	2003	FGD[2]	2008-2010	Electrostatic precipitator	pre-1990			
Wansley (7946)	2004	Natural Gas	Dry Low NO_x burners	2004	SCR[1]	2004							
Wansley CC (55965)	2012	Natural Gas	Dry Low NO_x burners	2012	SCR[1]	2012							
Yates	1950	Coal	Low NO_x burners (4/7 units)	1994 (2 units) 1995 (2 units)	Natural Gas Co-firing	1999	FGD[2] (1/7 units)	pre-1990	Electrostatic precipitator	pre-1990			

[1] Selective Catalytic Reduction
[2] Flue Gas Desulphurization

3.3.3 EGU policies enacted, 1990-2015

The EGUs included in this study are governed under both national and state rules. Major national rules implemented under the 1990 Clean Air Act (CAA) Amendments since 1990 include the 1990 Acid Rain Program (ARP), the 1998 NO_X SIP Call and associated Budget Trading Program (NBTP) and associated SIP Call (U.S. EPA, 2009), the 2008 Clean Air Interstate Rule (CAIR) (U.S. EPA, 2005), and the 2011 Cross-State Air Pollution Rule (CSAPR) (U.S. EPA, 2015). State rules are established through State Implementation Plans (SIPs), which are required by the EPA for all areas that are in non-attainment under the National Ambient Air Quality Standards, which are established under the CAA.

Many of the state rules align with and are driven by the national rules, but can be more specific, e.g., specifying certain emission levels or controls on specific plants. For instance, when many plants were required to implement seasonal controls under Georgia rules, many other eastern states were required to implement similar seasonal controls under the NBTP. Similarly, when the Georgia Multipollutant Rule was implemented beginning in 2009 that required stricter, year-round controls on NO_X and SO_2, the first phase of CAIR, which also required year-round controls was implemented on a national level. In general, the Georgia state rules reduce utilities' flexibility---both in timing and by dictating the specific plants that require controls---in that they require specific controls on specific plants that contribute to elevated air pollution concentrations in NAA's.

This interplay between national and state rules contributes to a blurring of lines between emissions reductions attributable to specific actions. On one hand, the utilities must keep their emissions across their fleet below the national standards. On the other, the state may require multiple plants that impact air quality in its NAAs to install specific controls by certain dates. By installing the required controls, the utility can claim emissions reductions under the national program while simultaneously adhering to the rules in the SIP.

The following details major programs of interest to the current thesis. The list is not exhaustive of all programs impacting EGU sources in the Southeast, but covers major regulations.

3.3.3.1 National rules

National Ambient Air Quality Standards (NAAQS). The EPA is required to set NAAQS in the United States under the CAA. These are revisited and revised periodically by the EPA. Attainment of these standards in certain areas is based on observations, and the EPA designates certain places non-attainment status. Each state that includes a NAA much submit a State Implementation Plan (SIP) to the EPA that details the state's plan to reduce ambient pollution levels to concentrations below the NAAQS. The EPA sets NAAQS for ozone, particulate matter (both $PM_{2.5}$ and PM_{10}), lead, SO_2, nitrogen dioxide (NO_2), and carbon monoxide (CO) (National Research Council, 2004).

The Acid Rain Program (ARP). The Acid Rain Program was enacted in 1993 to combat increasing SO_2 and NO_X concentrations throughout the United States, and

especially in the Eastern states. With the Clean Air Act Amendments on 1990, the EPA set out to reduce annual SO_2 emissions in the United States to 10 million tons less than they were in 1980 (U.S. EPA, 2002). To achieve these reductions, the EPA designed an approach that included two phases. Phase I, which began in 1995, targeted the largest existing power plants. Between 1990 and 1995, Georgia Power saw a dramatic decrease in SO_2 emissions (Ewald, 2015). Starting in 2000, Phase 2 required all other plants regulated under title IV of the CAA to achieve emissions reductions. To ensure reductions were being made, continuous emissions monitors were required for both SO_2 and NO_X on all regulated stacks (Harrington, et al., 2012).

The NO_X Budget Trading Program (NPB) and SIP Call. To address the problem of ozone precursors being transported across state lines in the East, the EPA issued the NO_X State Implementation Plan (SIP) Call in 1998. This Call was meant to improve the implementation of the controls established under the Acid Rain Program. The SIP Call did not place a limit on individual sources; instead, each state is required to reduce NO_X emissions during the ozone season to avoid non-attainment (U.S. EPA, 2009). The NO_X SIP Call only targeted coal plants during ozone season (5 months of the year) (Lloyd, 2014).

The EPA began the NBTP under the 1998 SIP Call to aid states in their effort to meet their emissions budgets. The NBTP was a cap-and-trade strategy that was optional; however, all 20 states and the District of Columbia used the program to help meet their NO_X SIPs by 2007. Georgia, which required seasonal NO_X controls on EGU sources in the NAA beginning in 2000, was not included under the NO_X SIP call or NBTP.

The Clean Air Interstate Rule (CAIR). CAIR, promulgated in 2005, was the regulatory approach to further reducing NO_X and SO_2 emissions adopted by the EPA after the Clean Skies Act did not pass Congress. The focus of CAIR is $PM_{2.5}$ transport across state borders. The regulation affected 28 eastern states, and set up three interstate emissions trading programs: the CAIR SO_2 annual trading program, the CAIR NO_X annual trading program, and the CAIR NO_X ozone season trading program. In effect, CAIR extended the ozone-season NO_X controls under the NPB to the entire year, and required large coal plants to install SO_2 controls or shut down. At http://www.epa.gov/airmarkets/programs/cair/ga.html (visited 3 September, 2015), the

EPA reports that CAIR will lead to reductions of NO_X and SO_2 emissions of 37% and 54%, respectively, in 2015.

A 2008 court decisions kept CAIR in place, but instructed the EPA to develop an alternative rule that satisfies CAA requirements related to cross-state transfer of air pollutants.

The Cross-State Air Pollution Rule (CSAPR). This rule was released by the EPA in July 2011 as the answer to the court ruling that reinstated CAIR. However, this rule was overturned by the courts before any parts of it were implemented, and CAIR remained in place. In October 2014, the D.C. Circuit issued a ruling that implementation of CSAPR would be delayed by three years, and the first phase replaced CAIR on 1 January, 2015.

3.3.3.2 State rules

The EPA requires states that contain NAAs to submit SIPs. These detail the state's plans for meeting NAAQS in the affected areas. In general, rules regarding EGUs in Georgia target only sources within or near NAAs, the largest of which is the Atlanta 20-County ozone and $PM_{2.5}$ NAA. The rules are documented in the Georgia Rules for Air Quality Control (GRAQC) and the SIPs submitted by the EPD to the EPA. In this section, reference will be made to both the GRAQC and the SIPs whenever possible. For rules in the GRAQC, we use the notation $GRAQC_{*}$ where * represents the rule section.

Emissions of Nitrogen Oxides from Major Sources ($GRAQC_{yy}$). Originally, the 13 counties of the ozone NAA region were included: Cherokee, Clayton, Cobb, Coweta, DeKalb, Douglas, Fayette, Forsyth, Fulton, Gwinnett, Henry, Paulding, and Rockdale. Major plants in these counties were McDonough and Yates. The rule requires that plants in these counties install 'reasonally available control technology' (RACT), as approved by the Director (of the EPD). The date of compliance was 31 July, 1995. In 1999, the RACT requirement was expanded to 32 more counties. Newly covered facilities were required to be in compliance.

In 2000, all facilities that fell under rules $GRAQC_{jjj}$, $GRAQC_{lll}$, $GRAQC_{mmm}$, $GRAQC_{nnn}$ (which included most major NOX sources within the NAA, and some outside of it) were exempt from rule $GRAQC_{yy}$, since the new rules were more stringent. In 2005, the rule was further amended to include smaller plants (with a compliance date of 1 May, 2007). The EPD added a public comment opportunity after each RACT approval.

NO_X Emissions from Electric Utility Steam Generating Units ($GRAQC_{jjj}$). This rule was adopted in 1998, and applied to EGUs within the 13-county NAA. The rule established summertime (1 May - 30 September) emissions limits on a lb $mmbtu^{-1}$ heat input basis. Compliance was required by some units as early as summer 1999. More plants were added each summer between 2000 and 2002. In 2000, the rule was amended to include Putnam County, home of Plant Branch.

This rule aligns with EPA's NBTP and SIP Call. Neither program was not implemented in Georgia, partly because this rule accomplished the same goal (U.S. EPA, 2008).

NO_X Emissions from Fuel-Burning Equipment ($GRAQC_{lll}$). This rule, adopted in 1999, sets an emissions limit of 30 ppm NO_X @ 3% O_2, dry basis on medium-sized sources during the summer months in 45 Georgia counties. Compliance was required beginning on 1 May, 2000.

NO_X Emissions from Stationary Gas Turbines and Stationary Engines used to Generate Electricity ($GRAQC_{mmm}$). Promulgated in 1999, this rule sets emissions limits on smaller plants (0.1 MWe to 25 MWe) beginning in May 2003.

NO_X Emissions from Large Stationary Gas Turbines ($GRAQC_{nnn}$). Promulgated in 2001, this rule sets emissions limits on large gas turbines, 45 counties. Old turbines (permitted before April 2000) were required to comply by May 2003.

Multipollutant Control for Electric Utility Steam Generating Units ($GRAQC_{sss}$). This rule was promulgated in 2007, and established dates of compliance that specific plants needed to install selective catalytic reduction (SCR) and flue gas desulfurization (FGD) on specific units. Further, the rule specifies dates of operation of the control systems (e.g. the SCR on Scherer is only required to be run in the summer), however, all plants in the NAA are required to run their controls year-round. The rule specifically mentions NO_X, SO_2, and mercury as the target pollutants of the rule.

This rule overlaps both in date and purpose with CAIR, in that it required year-round controls on NO_X, and strict controls on SO_2 on coal-fired power plants. However, the targeted approach on specific EGU units had the effect of reducing the overall flexibility of utilities to meet their limits under the CAIR cap-and-trade programs.

Mercury Emissions from New Electric Generating Units ($GRAQC_{ttt}$). Promulgated in 2007, the rule states that new units (permitted after 1 January, 2007) must install best available control technology (BACT) for mercury, as determined on case-by-case basis by the Director (of the Georgia EPD). The rule was later removed.

SO_2 Emissions from Electric Utility Steam Generating Units ($GRAQC_{uuu}$). This rule establishes SO_2 emissions limits on coal-fired power plants, including all of the plants in the current study plus Plants Hammond and Scherer. The dates of compliance correspond to the dates from the Multipollutant Control Rule ($GRAQC_{sss}$).

3.3.4 Attributing EGU emissions to specific policies

Reductions in NO_X emissions compared to the counterfactual (i.e., hypothetical scenario assuming various regulatory programs were not implemented) show several distinct periods in which daily emissions dropped in Atlanta (black dots in Figure 3-1). A value of zero on the graph represents no change from the counterfactual, and a value of -1.00 represents a 100% reduction compared to the counterfactual.

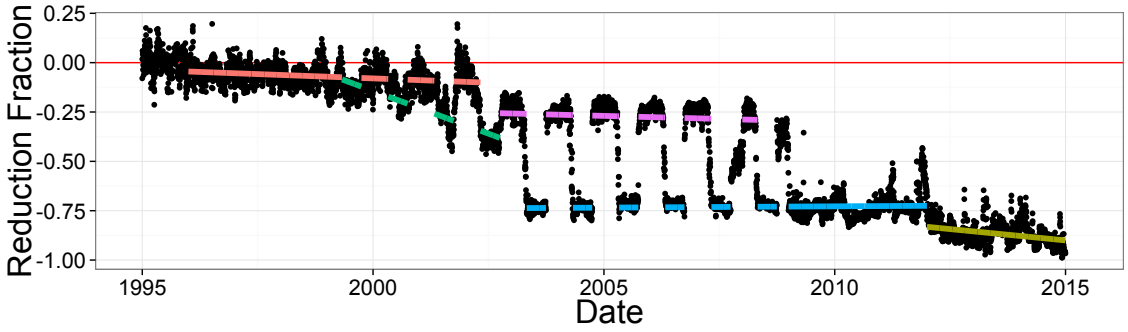


Figure 3-1. Fraction reduction in NO_X emissions normalized to the counterfactual (red line). Thicker lines correspond to time periods used to fit lines for quantifying reductions related to specific regulatory actions.

Five separate lines were fit to the graph in Figure 3-1. Each represents a different period (not always continuous) that is associated with a particular regulation. Lines indicate the time periods used to fit the lines, not the total period that the benefit of each control will be counted in the final analysis (this is discussed below). The lines were formed by fitting a linear regression to specific periods identified as dates of known control actions. Six dominating features of the graph are:

Orange (1996-2015). A 4.7% reduction is observed in the total average emissions factor in Atlanta for all EGUs during 1995. The line was only fit during the winters in the early 2000's because of the start of summertime NO_X controls in 1999.

Green (Summers 1999-2002). This period represents the beginning of the implementation of summer controls under $GRAQC_{jjj}$. The large negative slope of this line is due to the increasing number of units with SCR's at Bowen, Yates, and Wansley.

Purple (Winters 2002-2008). Under $GRAQC_{jjj}$, plants were required to control ozone during the ozone season. Since Plant Harllee Branch achieved this by installing low NO_X burners—which cannot be turned off like SCRs can—in 2002, the plant achieved emissions reductions year-round.

Blue (Summers 2003-2009, October 2009-December 2011). $GRAQC_{jjj}$, in full effect by summer 2003, required summertime NO_X controls. After 2008, $GRAQC_{sss}$ and CAIR required year-round NO_X controls, meaning the SCRs at Bowen, Yates, and Wansley were used both in the summer and the winter.

Yellow-green (2012-2014). Beginning in 2012, large portions of the load were generated with natural gas. A large portion of this was from Plant McDonough, although other coal plants were retrofitted during this period (Wansley), and existing smaller natural gas plants ran more as well.

There are a few sections in the graph in Figure 3-1 that do not match up with dates in the GRAQC. In the fall of both 2007 and 2008, it appears that some of the summertime controls remained active. 2008 is described by the EPA as a CAIR 'training year,' so it is possible that there is some connection here. Many of the deviations later in the time series (2010 and later) may be caused by startups and shutdowns and maintenance operations.

The linear approximations of known controls and specific dates from regulatory actions in Figure 3-1 were used to attribute emissions reductions to specific controls in Figure 3-2. The assumption was made that controls would continue to yield the same reduction in NO_X emissions after new regulations were put into place. This ensures that regulations are only attributed *additional* reductions on top of those realized under previous actions. A similar approach to that described above was applied to assess regional emissions (made of seven states—Alabama, Georgia, Mississippi, North Carolina, South Carolina, and Tennessee) reductions.

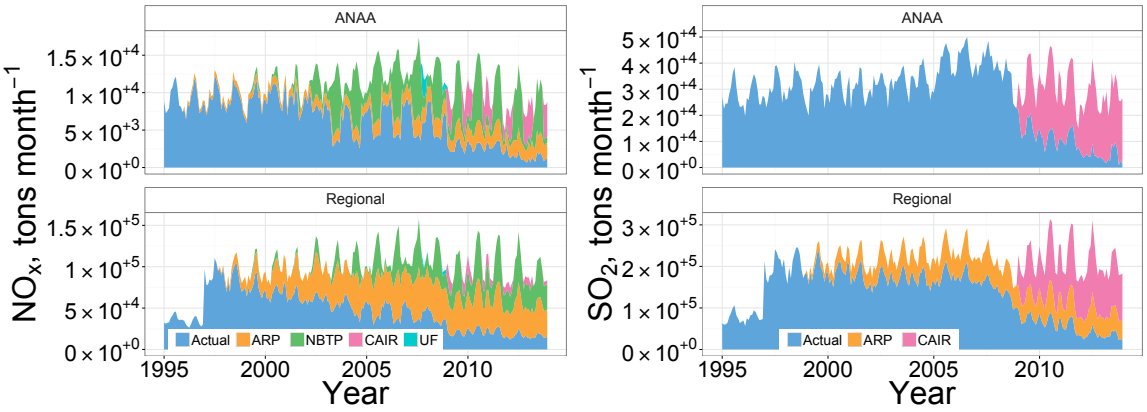


Figure 3-2. Changes in EGU emissions attributable to specific regulatory actions for NO_X (left) and SO_2 (right) in the ANAA (top) and Southeast Region (bottom). Regional CEM data was not completely reported until 1997, so all controls were assessed after this year. Labels: Actual—reported NO_X emissions, ARP—$GRAQC_{yy}$ and the ARP, NBTP–$GRAQC_{sss}$ and NBTP, CAIR—$GRAQC_{jjj}$ and CAIR, and UF—reductions that are not linked with specific controls. UF reductions may be related to test periods for controls that were not required by law. No UF periods were found for SO_2.

Actual NO_X emissions are measured at the source and reported by the EPA (U.S. EPA, 2016). Total NO_X emissions in the Atlanta area in 2014 were 89% less than in 1995 (Table 3-2).

Table 3-2. Emissions changes for various species and sources between reference year (y^*) and 2013.

Source (y^*)	Emissions in y^* (tons)	Emissions in 2013 (tons)	Percent change
EGU (1995)			
NO_X	303	43	86%
SO_2	920	142	85%
REG (1997)			
NO_X	2710	475	82%
SO_2	5604	943	83%
MOB (1993)			
NO_X	567	127	78%
SO_2	15	1	93%
$PM_{2.5}$	30	10	67%
CO	4306	1421	67%
VOC	326	123	62%
EC	12	4	61%
OC	18	7	67%

The first round of changes in Figure 3-1 and 2-2 aligns well with $GRAQC_{yy}$ and the Acid Rain Program. Many plants reduced their emissions slightly over the course of

1995, including Harllee Branch (2.5% reduction in emissions factor) and Wansley1 (2.6% reduction in emissions factor). Other coal plants reduced their emissions factor as well, resulting in a 4.7% reduction in the emissions factor across all plants. The negative slope of the corresponding fitted line in Figure 3-1 yields the increasing benefit of these regulatory programs over time.

The second group of reductions begins in 1999 with the onset of summertime NO_X controls. Beginning in Fall of 2003, the reductions include the year-round controls at Harllee Branch, since these are listed under the same rule ($GRAQC_{jjj}$). As discussed earlier, these reductions are attributed to both $GRAQC_{jjj}$ and the NBTP/SIP Call. Plant Bowen began operation of Selective Catalytic Reduction (SCR) control on 1/4 units in 1999, and completed the installation of SCRs on the three remaining units by 2003. Plant Wansley completed installation of its SCRs by 2003. These controls were only run in the summer months for the years 1999-2008.

The third major set of reductions began at the beginning of 2009 with the year-round controls required under $GRAQC_{sss}$ (The Multipollutant Control Rule) and CAIR. Because ozone-season controls were already in place, the benefits of these rules are restricted to the wintertime. Most the NO_X emissions reduction due to this rule comes from running the SCRs that were already installed at plants around Atlanta.

Around the time that the Multipollutant rule was beginning to take effect, natural gas prices in the United States began to fall (Figure 3-3). This, along with the price of the controls required under the multipollutant rule (for instance, under sections 7 and 8 of the rule, plant McDonough would have been required to install SCRs on and FGD on both units by May 2012), made the economics of converting much of the coal load to natural gas preferable to the alternative. (Figure 3-1, green line)

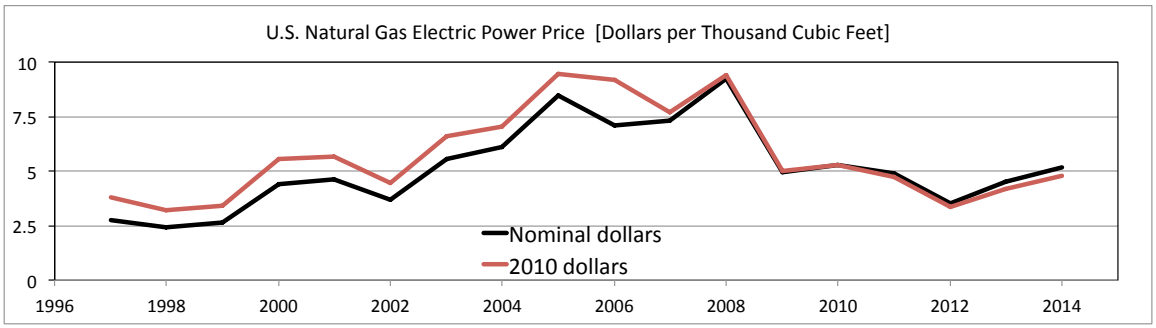


Figure 3-3. Price of natural gas in the United States. Source: U.S. Energy Information Administration. http://www.eia.gov/dnav/ng/hist/n3045us3a.htm Accessed 8 August, 2015.

The major retrofits to natural gas occurred at plants McDonough and Wansley. Further, many smaller natural gas that were constructed between 1999 and 2002—including Chattahoochee, Doyle, Hawk Road, MPC, Tenaska, and Walton—were online a larger percent of days in the later years compared to the years after they were first built. At the time these smaller plants first went online, they were mainly used in the summers when demand is highest.

Because a change to natural gas cannot be attributed solely to either regulatory actions or a reduction in the price of natural gas, the reductions achieved through a switch to natural gas are counted along with the $GRAQC_{jjj}$ and NBTP/SIP Call group in the summer and $GRAQC_{sss}$ and CAIR in the winter.

Uncategorized Controls. Two periods--the Falls of both 2007 and 2008--show lower than expected NO_X emissions based on the assessment of regulatory programs. It is possible that during these time periods one or more plants kept their SCR's turned on for one reason or another, but it is not immediately obvious why they would do this (since controls would have to be shown to be necessary to the PSC for cost-recovery).

SO_2 counterfactuals were developed in the same way as NO_X counterfactuals. Figure 3-4 shows the fraction of reduction in SO_2 emissions over time compared to the counterfactual. For most of the study period, the actual emissions stay near the counterfactual. Beginning in 2008, however, the installation of FGD controls at Bowen and Wansley reduced SO_2 emissions substantially. After 2010, the reduction fraction remained relatively constant until 2012, when the switch to large portions of the load being produced via natural gas caused the emissions to fall even further.

Because of the relationships between the Georgia Multipollutant Rule ($GRAQC_{sss}$), CAIR, and the emissions changes due to the economics of natural gas discussed above, all of the reductions in SO_2 seen in Figure 3-4 are attributed to $GRAQC_{sss}$ and CAIR.

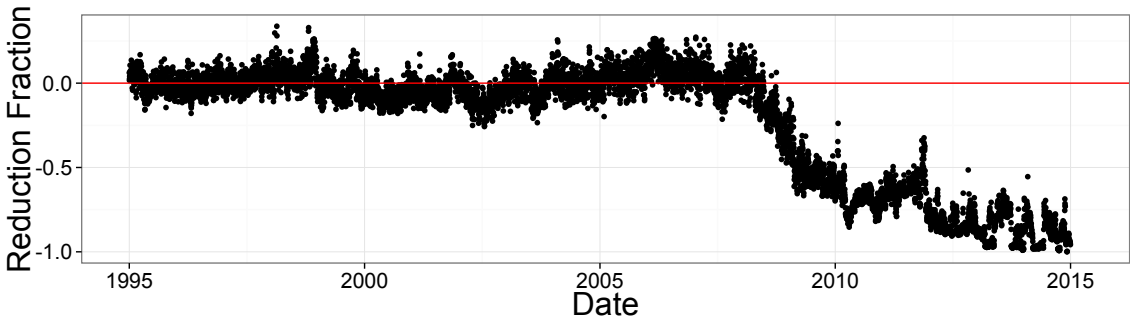


Figure 3-4. Fraction reduction in SO_2 emissions normalized to the counterfactual (red line).

3.4 Mobile source emissions

Mobile emissions are regulated to reduce emissions of air pollutants that are associated with adverse human health and environmental outcomes. Regulatory programs are implemented by national (e.g. the United States Environment Protection Agency, EPA) and state/local (e.g. the Georgia Environmental Protection Division, EPD) regulating agencies.

Because of the costs associated with the implementation of controls installed to comply with these regulations, there is interest in assessing the effectiveness of the regulations. Previous studies have undertaken this type of accountability research on regulatory programs implemented under the 1990 Clean Air Act Amendments (Harrington et al., 2012; Morgenstern et al., 2012; van Erp et al., 2008 & 2011). This paper seeks to extend previous research by investigating four specific programs in Atlanta: the Enhanced Vehicle Inspection and Maintenance Program (1996-present), the Georgia Gasoline Marketing Rule (1999-present), Tier 2 Motor Vehicle Emissions Standards and Gasoline Sulfur Requirement (2006-present), and the Heavy-Duty Highway Rule (2004-present).

Enhanced inspection and maintenance has been required on automobiles registered in 13 counties surrounding Atlanta since October 1996. The affected counties are Cherokee, Clayton, Cobb, Coweta, DeKalb, Douglas, Fayette, Forsyth, Fulton, Gwinnett, Henry, Paulding, and Rockdale. In general, the requirement covers gasoline-powered cars and light trucks, specifically 24 model years old and newer.

Two separate gasoline programs are included in this analysis. Beginning in 1999, Georgia required gasoline sold in an expanded region of 25 counties to have a Reid Vapor Pressure (RVP) of 7 psi or less and a seasonal average sulfur content of less than 150 ppm (by weight). The seasonal sulfur limit was reduced in 2003 to 90 ppm and in 2004 reduced further to a 30 ppm year-round average. In 2006, the federal Tier 2 limit required an annual average of 30 ppm or less. Therefore, before 2006, benefits are attributed to Georgia Gasoline and after to the Tier 2 program.

Lowering RVP is an approach to reduce evaporative VOC emissions. Reducing sulfur both reduces SO_2 and sulfate emissions and improves the efficiency of selective catalytic reduction controls that reduce NO_x emissions. The focus of both rules—the Georgia Gasoline Marketing Rule and the Tier 2 Motor Vehicle Emissions Standards and Gasoline Sulfur Requirement—was to reduce mobile contributions to ambient ozone and PM levels by reducing NO_X and VOC emissions. The rule has the further effect of reducing SO_X emissions, which contribute to formation of secondary PM through atmospheric conversion to SO_4^{2-} (U.S. EPA, 1999).

The Tier 2 program included updated engine emissions standards applicable to all passenger cars, light trucks, and medium-duty passenger vehicles. The standards were phased in between 2004 and 2009. New vehicles under this program are required to average 0.07 g mi^{-1} NO_X, and non-methane organic gases (NMOG–a component of VOCs) are regulated based on which of several bins each car fits into. NO_X standards before this rule was in place ranged from 0.30 g mi^{-1} to 1.53 g/mi depending on the type of vehicle.

The 2007 Heavy Duty Highway Rule was promulgated in 2001. Like the Tier 2 gasoline rule, this program sets standards for both engines and fuel. The goal of this legislation was to reduce ozone levels by reducing ozone precursor emissions (NO_X and non-methane hydrocarbons–NMHCs, a component of VOCs). One major aspect of the rule was limiting sulfur content to 15 ppm or less by June 2006. According to information from the Energy Information Administration, diesel sales in Georgia went from being comprised of 91% fuel with sulfur content between 15 and 500 ppm in 2006 to 35% in 2007, with the difference being declining sales of diesel with sulfur content greater than 500 ppm and increasing sales of fuel with sulfur content less than 15 ppm. By 2008, no diesel with sulfur

content greater than 500 ppm was sold, and by 2012, 100% of fuel sold in Georgia had less than 15 ppm sulfur (U.S. Energy Information Administration, 2014).

A second major component of the Heavy Duty Highway Rule is the reduction of NO_X and NMHC emissions standards applicable to all highway heavy-duty engines. These standards were enforced beginning on model year 2004 vehicles. The previous standards were (for all hydrocarbons–HCs) 1.3 g bhp-hr^{-1} and (for NO_X) 4.0 g bhp-hr^{-1}. The updated standard is a combined limit of 2.4 g bhp-hr^{-1} for combined NO_x and NMHC or 2.5 g/bhp-hr with a limit of 0.5 g bhp-hr^{-1} on NMHC (U.S. EPA, 1997). In 2007, standards were introduced to be phased in 2007-10 that limited NMHC emissions to 0.14 g bhp-hr^{-1} and NO_x emissions to 0.2 g bhp-hr^{-1}.

3.4.1 Mobile emission estimates

Daily mobile emissions for 2000-2012 in the former 1997 ozone National Ambient Air Quality Standards (NAAQS) 20-county Atlanta Non-Attainment Area (NAA–Barrow, Bartow, Carroll, Cherokee, Cobb, Clayton, Coweta, DeKalb, Douglas, Fayette, Forsyth, Fulton, Gwinnett, Hall, Henry, Newton, Paulding, Rockdale, Spalding, and Walton) were estimated using the EPA's MOtor Vehicle Emissions Simulator (MOVES) version 2010b (U.S. EPA, 2012b). National default data was supplemented with locally developed or interpolated data when available. Local data included monthly average meteorology from JST, inspection and maintenance records from the Georgia EPD, and Average Annual Vehicle Miles Traveled (AAVMT) from the Georgia Department of Transportation web database (http://www.dot.ga.gov/DS/Data#tab-2; "445 Reports (Mileage By Route Type and Functional Classification)" in the "400 Series Reports" drop-down box). VMT was distributed by road type by the Atlanta Regional Council's Travel Demand Model. Local vehicle population and age distribution for 2002 were purchased by EPD from R. L. Polk & Company and extrapolated to future years. Fuel formulations are MOVES defaults that have been applied based on the specifications applied to each year in this modeling. Changes in the population for Fulton County, which contains a majority of the Atlanta city area, were used to grow vehicle populations.

3.4.1.1 Modeling estimated and counterfactual emissions

Raw daily emissions estimates from MOVES show nonrealistic features due to specific parameters in MOVES, such as fuel types used in different months, changes in

AAVMT, and annual changes in fleet distributions. To smooth out these features, emissions from each pollutant were estimating using the best fit of a linear model with the following covariates: time, time squared, weekday/weekend factor, sine and cosine with period of 1 year, and time multiplied by cosine with period of 1 year. The result is a smooth time trend that maintains long-term changes, annual fluctuations, and weekend-weekday variability for each species (Figure 3-5).

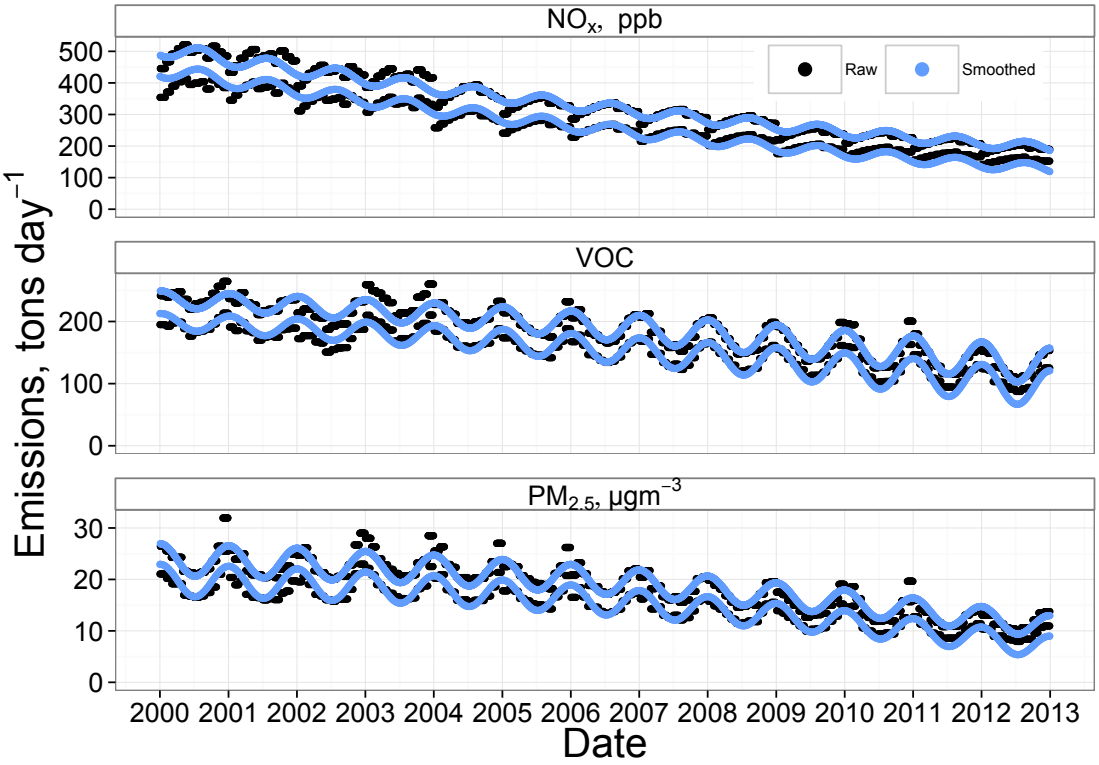


Figure 3-5. Raw MOVES-modeled and smoother mobile emissions estimates. Monthly and annual variability in MOVES estimates is generated by changes in fuel type, meteorology, and default VMT by month. Weekend-weekday differences appear as the two parallel lines (weekend is the lower of the two).

Mobile sources have unique features that complicate the estimation of counterfactuals. First, the ANAA fleet is made up of vehicles of varying ages, manufacturers, types, and fuel types that are subject to different regulations. The counterfactual mobile emissions scenarios required alternative specifications within and outside of MOVES to estimate counterfactual emissions. IM scenario emissions were estimated by clearing the inspection and maintenance table from the input database—functionally, this equates to eliminating the IM program—and rerunning MOVES for the

entire period of interest. IM was required beginning in 1993 in a 13-county subsection of the ANAA; the model was rerun for this subset.

The other scenarios required a different approach, and the following explanation requires a clarification of the terms. Model Year (MY) is associated with when the vehicle first entered the vehicle population, and emission year (EY) is the year of interest. Emissions factors for automobiles of a certain MY generally increase with each passing EY at varying rates depending on the vehicle type, fuel type, and other factors according to models within MOVES. Because of updates to automobile engines, changes in fuel composition, and other changes, cars with later MYs tend to have lower emissions factors as well.

For the all counterfactual scenarios, which assume no new mobile source regulations after 1993, emissions factors for EY 1993 were applied to all future years. The emissions factors were assigned by pollutant, fuel type, process type (e.g. running exhaust, refueling displacement vapor loss, etc.), source type (e.g. passenger car, motorcycle, transit bus, etc.), month, day (weekend or weekday), and MY. For full lists of each segment, see U.S. EPA 2012. In EY 2000, for example, a MY 1998 gasoline-powered passenger car is assigned the complimentary emissions factor for a MY 1991 gasoline-powered passenger car in EY 1993. This approach was corroborated by running MOVES with the 'Rate of Progress' option, which models a scenario with no Clean Air Act Amendments by applying 1993 emission rates to all vehicles after this year. Results from the two approaches were identical.

The approach for the gasoline and diesel programs (GSP and DSP) is similar; the above method for applying emissions factors for previous EY's was applied to only vehicles that used the fuel type of interest. In the GSP scenario, emissions factors for EY 1999 were applied for all future years, and in the DSP scenario, emissions factors for EY 2005 were applied for future years.

Results show large reductions in emissions of most species over the period of interest (Figure 3-6). Estimated changes in emissions due to each control show that the Tier 2 gasoline programs program led to the greatest emissions reductions in both NO_X and VOCs. Modeled ANAA mobile NO_X, VOC, CO, and $PM_{2.5}$ emissions decreased by 78%, 62%, 67%, and 66% respectively between 1993 and 2013 (Table 3-2).

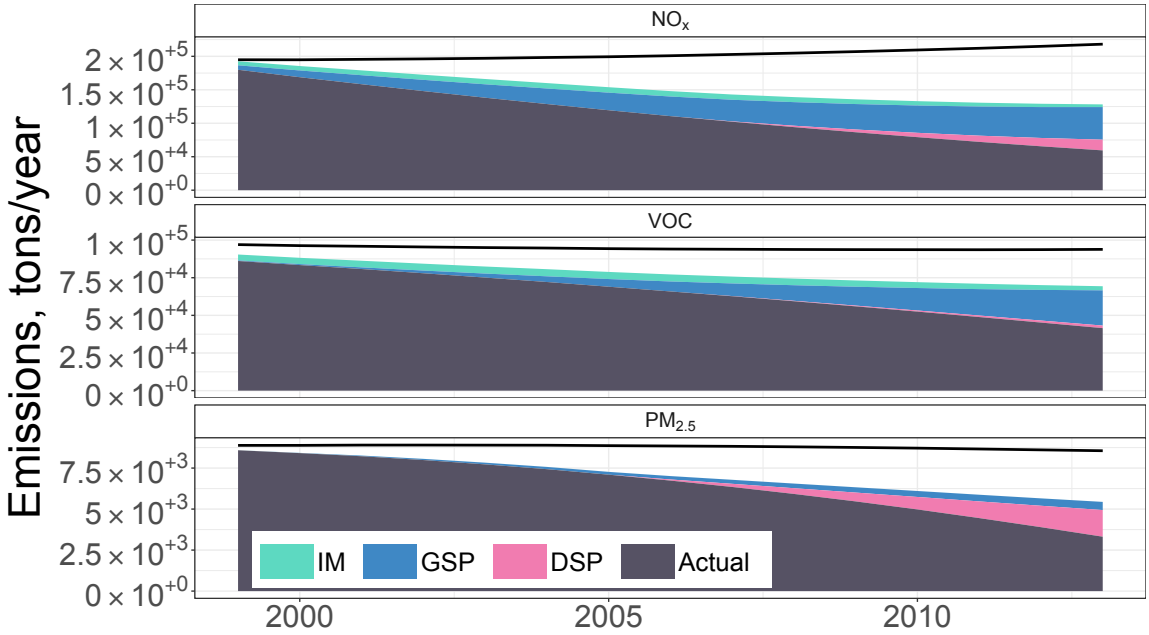


Figure 3-6. Actual emissions and emissions avoided by each program: Enhanced Vehicle Inspection and Maintenance (IM), the Georgia Gasoline Marketing Rule and Tier 2 Motor Vehicle Emissions Standards and Gasoline Sulfur Requirement (GSP), and the Heavy-Duty Highway Rule (DSP). The black line at the top of each plot represents the 'all counterfactuals' scenario. The difference between 'all counterfactuals' scenario and the reductions attributed to other controls is due to vehicles not covered by the control programs, e.g., gasoline vehicles with model year before 2000.

Of the three specific programs investigated, IM programs had the smallest effect on emissions across all years. The impact of IM controls diminished slightly over the time period of interest. Gasoline programs had relatively larger impacts on NO_X, VOC, and CO, while diesel programs impacted $PM_{2.5}$ emissions the most. Benefits from the diesel programs did not begin until after 2006, when the diesel rule came into effect.

The sum of the mobile policies does not describe the emissions changes estimated in the MOB_{ALL} scenario because of other policies that were implemented between 1993 and 2013 that are not being examined here. These include the standards established under the 1990 Clean Air Amendments themselves (https://www.epa.gov/clean-air-act-overview/1990-clean-air-act-amendment-summary-title-ii) and the Low Emissions Vehicle Program (https://www3.epa.gov/otaq/lev-nlev.htm#implementation) implemented in the late 1990s. There are other reasons car makers would change engines that may not be regulation-driven as well, e.g., performance measures.

3.5 Conclusions

This chapter presented an in-depth review of air quality regulations that have altered emissions of major sources in the southeastern United States. The result of this chapter is multiple counterfactual emission time series of multiple pollutants from EGU and on-road mobile sources. Differences between actual and counterfactual emissions are attributable to air quality regulations; however, utility operators and auto makers make decisions based on multiple factors, so it is impossible to precisely separate the impacts of various individual regulatory programs.

This portion of the analysis is important in an accountability framework; indeed, the counterfactuals developed here inform much of the work that follows in this thesis. While the list is not exhaustive, the programs discussed here are the most important in terms of wide-spread emissions reductions. Previous accountability studies have failed to relate regulations to emissions changes, and have instead attempted to relate interventions directly to changes in air quality or health. This approach invites potential confounding factors, such as concurrent emission changes or improvements in efficiency, into the analysis. The current approach, while limited by assumptions in the emissions models, alleviates many of the confounding factors inherent in previous studies by linking changes in emission factors to implementation dates of specific control policies. Discussions with stakeholders and comparisons of results with known control implementation dates provides evidence that the counterfactuals capture regulatory-relevant impacts, and avoids influence from other factors.

3.6 Acknowledgements

This material is based upon work supported by Health Effects Institute and the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1148903. I am especially appreciative of Charles Huling, Steven Ewald, and David Lloyd for conversations regarding power plant regulations.

3.7 References

Ewald, S. (2015). Private Communication.

Harrington, W., Morgenstern, R., Shih, J.-S., & Bell, M. L. (2012). Did the Clean Air Act Amendments of 1990 really improve air quality? *Air Quality, Atmosphere & Health*, *5*(4), 353–367. https://doi.org/10.1007/s11869-012-0176-5

Health Effects Institute. (2003). *Assessing Health Impact of Air Quality Regulations: Concepts and Methods for Accountability Research*. Retrieved from http://pubs.healtheffects.org/getfile.php?u=261

Huling, C. (2014). Private Communication.

Lloyd, D. (2014). Private Communication.

Muller, N., Tong, D., & Mendelsohn, R. O. (2009). Regulating NOx and SO2 Emissions in Atlanta. *BE Journal of Economic Analysis & Policy: Contributions to Economic Analysis & Policy*, *9*(2). Retrieved from http://ideas.repec.org/a/bpj/bejeap/v9y2009i2n3.html

National Research Council. (2004). *Air Quality Management in the United States*. National Academies Press. https://doi.org/10.17226/10728

U.S. Energy Information Administration. (2014). *Georgia Prime Supplier Sales Volumes of Petroleum Products*. Retrieved from http://www.eia.gov/dnav/pet/pet_cons_prim_dcu_sga_a.htm

U.S. EPA. (1997). Heavy-Duty Highway Compression-Ignition Engines and Urban Buses–- Exhaust Emissions Standards. Retrieved January 1, 2015, from http://www.epa.gov/otaq/standards/heavy-duty/hdci-exhaust.htm

U.S. EPA. (1999). *Regulatory Impact Analysis–Control of Air Pollution from New Motor Vehicles: Tier 2 Motor Vehicle Emissions Standards and Gasoline Sulfur Control Requirements*. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Regulatory+Impact+Analysis+-+Control+of+Air+Pollution+from+New+Motor+Vehicles+:+Tier+2+Motor+Vehicle+Emissions+Standards+and+Gasoline+Sulfur+Control+Requirements#0

U.S. EPA. (2002). *Acid Rain Program 2001 Summary Report*. https://doi.org/10.1016/B978-0-08-091680-4.00015-9

U.S. EPA. (2005). *Regulatory Impact Analysis for the Final Clean Air Interstate Rule*. Retrieved from http://www.epa.gov/cair/pdfs/finaltech08.pdf

U.S. EPA. (2008). *Final Rule to Remove Georgia from the NOx SIP Call Requirements*.

U.S. EPA. (2009). *NOx Budget Trading Program - Basic Information*. Retrieved from http://www.epa.gov/airmarkets/progsregs/nox/docs/NBPbasicinfo.pdf

U.S. EPA. (2012a). *Motor Vehicle Emission Simulator (MOVES): User Guide Version, MOVES2010b*. Retrieved from https://www3.epa.gov/otaq/models/moves/documents/420b12001b.pdf

U.S. EPA. (2012b). MOtor Vehicle Emissions Simulator (MOVES). Retrieved from http://www.epa.gov/otaq/models/moves

U.S. EPA. Cross-State Air Pollution Rule Update for the 2008 Ozone NAAQS (2015).

U.S. EPA. (2016). Air Markets Program Data. Retrieved from https://ampd.epa.gov/ampd/

CHAPTER 4. METEOROLOGICAL DETRENDING OF PRIMARY AND SECONDARY POLLUTANT CONCENTRATIONS: METHODS APPLICATION AND EVALUATION USING LONG-TERM (2000-2012) DETAILED DATA[3]

4.1 Abstract

The effectiveness of air pollution regulations and controls are evaluated based on measured air pollutant concentrations. Air pollution levels, however, are highly sensitive to both emissions and meteorological fluctuations. Therefore, an assessment of the change in air pollutant levels due to emissions controls must account for these meteorological fluctuations. Two empirical methods to quantify the impact of meteorology on pollutant levels are discussed and applied to the 13-year time period between 2000 and 2012 in Atlanta, GA. The methods employ Kolmogorov-Zurbenko filters and linear regressions to detrended pollutant signals into long-term, seasonal, weekly, short-term, and white-noise components. The methods differ in how changes in weekly and holiday emissions are accounted for. Both can provide meteorological adjustments on a daily basis for future use in acute health analyses.

The meteorological impact on daily signals of ozone, NO_X, CO, SO_2, $PM_{2.5}$, and PM species are quantified. Analyses show that the substantial decreases in seasonal averages of NO_X and SO_2 correspond with controls implemented in the metropolitan Atlanta area. Detrending allows for the impacts of some controls to be observed with averaging times of as little as 3 months. Annual average concentrations of NO_X, SO_2, and CO have all fallen by at least 50% since 2000. Reductions in NO_X levels, however, do not lead to uniform reductions in ozone. While average detrended summer average maximum

[3] A modified version of this chapter has been published as: Henneman, L.R.F., Holmes, H., Mulholland, J., Russell, A. (2015) "Meteorological detrending of primary and secondary pollutant concentrations: Method application and evaluation using long-term (2000–2012) data in Atlanta, Atmospheric Environment, Volume 119, Pages 201-210, ISSN 1352-2310. Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.atmosenv.2015.08.007.

daily average 8h ozone (MDA8h O_3) levels fell by 4% (2.2 ± 2 ppb) between 2000 and 2012, winter averages have increased by 12% (3.8 ± 1.4 ppb), providing further evidence that high ozone levels are NO_X-limited and lower ozone concentrations are NO_X-inhibited. High ozone days (with MDA8h O_3 greater than 60 ppb) decreased both in number and in magnitude over the study period.

4.2 Introduction

Policy makers design regulations to reduce the adverse impacts of air pollutant emissions on human health and the environment. The effectiveness of the regulations are assessed based on measured pollutant concentrations and estimates of impacts on health endpoints related to exposure to these pollutant concentrations (van Erp et al., 2008). However, given the number of factors that affect a measured air pollution signal (e.g. meteorological fluctuations, climate change, population growth, weekend/weekday emissions differences, and the disproportionate effect of sources close to the monitor), it is often difficult to identify and quantify the effects of specific controls in long-term air pollution records. The meteorological impact on ozone and particulate matter with diameter less than 2.5 μm ($PM_{2.5}$) levels is especially important because of their known associations with negative health outcomes and the fact that ambient concentrations of both ozone and certain components of $PM_{2.5}$ are secondary, i.e., their formation in the atmosphere is dependent on both meteorological and chemical processes. Although emissions of precursors to ozone (e.g. NO_X and volatile organic compounds—VOCs) and $PM_{2.5}$ (e.g. primary $PM_{2.5}$, NO_X, SO_4^{2-}, NH_3, and VOC's) have decreased under programs implemented under the Clean Air Act since 1970, ozone and $PM_{2.5}$ levels continue to exceed EPA standards (Pachon et al., 2012; Vijayaraghavan et al., 2014). In order to assess the impact of controls on measured pollutant signals, it is important to account for the impact of variations in meteorology (Cox and Chu, 1993; Rao and Zurbenko, 1994).

The years between 2000 and 2012 saw the implementation of a number of regulations and associated controls on emissions across the United States. Major programs implemented during this period include the Tier 2 Vehicle and Gasoline Sulfur Program (phased in 2004-2009), the 2007 Heavy-Duty Highway Rule, the Acid Rain Program and associated NO_X State Implementation Plan (SIP) Call and Budget Trading Programs

(phased in 1995e2010), and the Clean Air Interstate Rule (CAIR phased in beginning in 2009) (Hubbell et al., 2009; Morgenstern et al., 2012; USEPA, 1999; USEPA, 2000; USEPA, 2009). In order to better assess the effectiveness of these regulations it is necessary to investigate the primary pollutants whose emissions are affected by the controls directly, the secondary pollutants that are products of atmospheric reactions of the primary pollutants (e.g. O_3 and $PM_{2.5}$), and the role that meteorological fluctuations play in affecting trends. To this end, meteorological effects can be quantified and removed from the measured pollutant signal.

A number of methods have been used to remove meteorological impacts on ambient pollutant concentrations (Cox and Chu, 1993; Flaum et al., 1996; Gardner and Dorling, 2000; Kuebler et al., 2001; Rao and Zurbenko, 1994). Rao and Zurbenko (1994) used multiple-pass moving average filters of different averaging windows and varying numbers of passes to isolate short-term fluctuations in ozone concentrations, and found a lag between temperature and ozone. Flaum et al. (1996) extended the analysis further to include dew point temperature, specific humidity, and wind speed. Kuebler et al. (2001) employed a filter and regression technique similar to that of Rao and Zurbenko (1994), but also included effects of solar radiation.

Others have used approaches besides filtering and regression techniques to isolate meteorological contributions. Cox and Chu (1993) used data from 43 urban areas to create a distribution of ozone concentrations in urban areas, and accounted for both surface temperature and wind speed effects. Gardner and Dorling (2000) used neural networks and filtering to calculate meteorology-adjusted ozone time series. In general, the neural network method finds a greater contribution of meteorology to ozone concentrations than the filter-regression methods. Brönnimann and Neu (1997) used weekly emissions trends derived from the Swiss emissions inventory and characterized meteorological conditions as “favorable” and “unfavorable” in order to determine concentration levels that can be expected on different days given certain meteorological conditions.

Camalier et al. (2007) used generalized linear models (GLMs) to quantify annualized effects of meteorology on ozone levels in cities across the United States. This method was extended by Blanchard et al. (2010) to investigate the effects of changing levels of ozone precursors on ozone levels in Atlanta and the reason for the relatively small

change in ozone levels compared to its precursors. The authors found that daily ozone levels are more sensitive to meteorological fluctuations than to changes in ambient non-methane organic carbon and NO_Y.

Within the context of an accountability assessment of how regulatory policies affect acute (daily) health outcomes, it is important to estimate meteorological impacts on observed daily concentrations for use in acute health impact analyses. The methods presented here assess more than a decade of observations of multiple gaseous and particulate constituents to aid the assessment of regulatory actions on public health, independent of daily meteorological fluctuations.

This paper presents results obtained using meteorological detrending methods used by Kuebler et al. (2001) and a new method designed to improve the ability to isolate only the fluctuations attributed to meteorological variability in a daily concentration signal. The new method uses multiple Kolmogorov-Zurbenko filters to account for the long-term and annual fluctuations, and a multiple linear regression with meteorological terms, interaction terms, monthly terms, and weekday and holiday terms. The methods are applied to ozone, NO_X, SO_2, CO, total $PM_{2.5}$, SO_4^{2-}, organic carbon (OC) and elemental carbon (EC) measured in Atlanta, GA from 2000 to 2012.

4.3 Data

Air quality and meteorological data, excluding rainfall, are from the SEARCH network's Jefferson Street monitoring station (JST) in downtown Atlanta (Atmospheric Research and Analysis, 2014). Rainfall measurements come from NOAA's National Climatic Data Center station at Hartsfield-Jackson Airport, which is 15 km south of the Jefferson Street Station.

Eleven pollutant species and eight meteorological metrics are used in this study (Table 4-1 Table 4-2). The meteorological variables are chosen as those likely having a strong physical-chemical link to the pollutant species. Daily metrics of three gas phase species (ozone, NO_X, and CO) are calculated in two ways each—the daily 8-h and 1-h maxima ozone and daily and morning means of NO_X and CO. Morning means capture commute emissions, while daily means reduce the impact of high concentrations seen due to shallow boundary layer height expected in the early morning (Kuebler et al., 2001). Most meteorological and species records had less than 10% missing days with the exceptions of

OC and EC, which switched from being measured every day to 1-in-3 day measurements in 2009. Hourly measurements of incoming solar radiation, temperature, wind speed, and relative humidity were converted to daily metrics (Table 4-2). Rainfall was used as a binary input.

Table 4-1. Daily pollutant species used in detrending analysis (2000-2012). Hourly measurements from JST are converted to daily metrics (# days = 4745).

Species[a]	Metric	Period	Exclusion criteria[b] (hours)	# Days	% Missing Days
$MD8hO_3$	Max of 8-h mean	12 a.m.-11 p.m.	—	4681	1.48
O_3^M	Daily max	12 a.m.-11 p.m.	≥ 12/24	4627	2.57
NO_X	Daily mean	11 a.m.-7 p.m.	≥ 5/9	4447	6.36
NO_X^{morn}	Morning mean	8 a.m.-11 a.m.	≥ 3/4	4398	7.39
CO	Daily mean	11 a.m.-7 p.m.	≥ 5/9	4609	2.95
CO^{morn}	Morning mean	8 a.m.-11 p.m.	≥ 3/4	4623	2.65
SO_2	Daily max	12 a.m.-11 p.m.	≥ 12/24	4658	1.92
$PM_{2.5}$	Daily mean	12 a.m.-11 p.m.	—	4015	9.41
SO_4^{2-}	Daily mean	12 a.m.-11 p.m.	≥ 12/24	4015	1.92
EC	Daily mean	12 a.m.-11 p.m.	—	3384	28.7
OC	Daily mean	12 a.m.-11 p.m.	—	3384	28.7

[a] The superscript M for pollutant abbreviations in the left column signifies the daily maximum, the superscript morn signifies a morning mean, and metrics with no superscripts are the average of hourly values within the Period.

[b] If the Exclusion Criterion for each averaging/maximum period is violated (i.e. there are more hours missing than allowed), the day is counted as NA. If there is no Exclusion Criterion, the minimum number of measurements needed for that day is one.

Table 4-2. Daily weather variables used in detrending analysis (2000-2012). Hourly measurements from JST are then converted to daily metrics (# days = 4745). Rainfall (RF) is the exception: measurements are from NOAA's National Climatic Data Center station at Hartsfield-Jackson Airport, which is 10 miles south of JST.

Phase	Metric	Period	Exclusion criteria (hours)	# Days	% Missing Days
SR	Daily total	12 a.m.-11 p.m.	Any, 7 a.m.-6 p.m.[c]	4950	9.29
SR^M	Daily max	12 a.m.-11 p.m.	—	5398	1.11
T^m	Daily mean	11 a.m.-3 p.m.	≥ 3/5	5368	1.66
T^M	Daily max	12 a.m.-11 p.m.	—	5425	0.62
WS	Daily mean	11 a.m.-3 p.m.	≥ 3/5	5284	3.19
WS^{morn}	Morning mean	7 a.m.-10 a.m.	≥ 3/5	5225	4.27
RH	Morning mean	8 a.m.-11 a.m.	≥ 3/4	5363	1.75
RF	Daily factor	12 a.m.-11 p.m.	—	5459	0

[a] SR—solar radiation, T—temperature, WS—wind speed, RH—relative humidity, and RF—rainfall. Superscripts *M* and *morn* signify the daily maximum and morning mean. Metrics with no superscripts are the average of hourly values within the Period.

[b] If the Exclusion Criterion for each averaging/maximum period is violated (i.e. there are more hours missing than allowed), the day is counted as NA. If there is no Exclusion Criterion, the minimum number of measurements needed for that day is one.

[c] If any solar radiation values are missing during the daytime hours, the day is excluded.

4.4 Methods

Observed pollutant species concentrations are a nonlinear function f^* of emissions (E) and meteorology (M), i.e.

$$C(t) = f^*[E(t), M(t)] \quad \text{(Eq. 4.1)}$$

Results from Blanchard et al. (2010) suggest a near-linear relationship between ozone and temperature and relative humidity, and nonlinear relationships for other meteorological variables (Blanchard et al., 2010). Recognizing that most variability in pollutant concentrations is on a seasonal scale, and emissions and their effects on measured pollutant species are primarily cyclical (dominated by daily, weekly, and annual patterns), their contributions to a daily signal can be removed along with the background. The residual fluctuations in the signal can then be attributed to short-term meteorological fluctuations and white noise through a multiple linear regression. Short-term meteorological fluctuations are assumed to be unaffected by and independent of emissions.

The long-term time series of daily pollutant concentrations (C) can be decomposed into components (Kuebler et al., 2001; Rao and Zurbenko, 1994):

$$\ln[C(t)] = C^{LT}(t) + C^{S}(t) + C^{W}(t) + C^{STM}(t) + C^{WN}(t) \quad \text{(Eq. 4.2)}$$

or

$$C(t) = \exp\left[C^{LT}(t) + C^{S}(t) + C^{W}(t) + C^{STM}(t) + C^{WN}(t)\right] \quad \text{(Eq. 4.3)}$$

Ozone concentrations in the atmosphere typically approximate a log-normal distribution (Hogrefe and Rao, 2000; Rao et al., 1997). A log transformation is applied prior to detrending for all pollutants to ensure model residuals follow the assumptions of normal distribution and homoscedasticity. A discussion of model residuals and comparing log- and no-log models is included in the Supporting information.

The components of each signal in (Eq. 4.2) are defined as follows (time period in parentheses):

LT: long-term (>365 days)

S: seasonal (365 days)

WH: weekly-holiday (7-365 days)

STM: short-term meteorological (<365 days)

WN: white-noise (1 day)

The LT component in the species decomposition captures long-term changes in background concentrations, climate, and precursor emissions (in secondary pollutants) (Gardner and Dorling, 2000; Rao and Zurbenko, 1994). While the WH, STM, and WN signals cover overlapping periods, they are each independent. The goal of the detrending is to quantify and remove the STM signal and relate it to meteorological fluctuations so that the remaining daily signal is corrected for short term meteorological fluctuations, i.e., deviations from the typical meteorology for the time of year.

A meteorological signal can be decomposed in a similar way as a pollutant species signal, though a log transform is not used for the meteorological variables. A log transform is not necessary because meteorological variables are the covariates in regression models described below, and there are no assumptions of specific distributions that need to be met. The short-term deviations of meteorological signals, ΔM, occur on time periods of less than those accounted for by the long-term and seasonal trends.

$$M(t) = LT^{met}(t) + S^{met}(t) + \Delta M(t) \quad \text{(Eq. 4.4)}$$

Two methods are assessed for their ability to separate daily deviations from long-term fluctuations. Both methods quantify the effect of individual meteorological variables on species signals. They are designed to create daily adjustments for meteorological

contributions to pollutant signals, and account for holiday and weekday effects that vary throughout the year as well as nonlinear effects of temperature and relative humidity on daily pollutant concentrations. While the results reported here are for data from Atlanta, the methods can be applied to any monitoring station with a multi-year meteorological and ambient air pollution record.

4.4.1 Multiple KZ and Kuebler et al. (2001) methods

The Multiple KZ (MKZ) and Kuebler et al. (2001) (KEA) methods are approaches to performing the detrending described above. Both use the Kolmogorov-Zurbenko (KZ) filter, which is a low-pass moving average filter (Kuebler et al., 2001; Rao and Zurbenko, 1994; Zurbenko, 1991). Zurbenko (1991) defined a $KZ_{m,p}$ as

$$Y_i = \frac{1}{m}\sum_{j=-k}^{k} X_{i+j} \qquad \text{(Eq. 4.5)}$$

where $m = 2k + 1$ (Zurbenko, 1991). m defines the window length, p is the number of passes, and i is the day of interest. KZ filters are widely used for their adaptability to different cutoff frequencies and robustness to missing data (Eskridge et al., 1997). Rao and Zurbenko (Rao and Zurbenko, 1994) investigated different values of m and p depending on the desired cutoff frequency. The filtering in this work was done using version 3.0.0 of the 'kza' filter package in the statistical program R (Close and Zurbenko, 2011; R Core Team, 2012). A benefit of the KZ filter as implemented in R is that it can handle missing data.

Both the MKZ and KEA methods employ KZ filters and linear regressions to separate the long-term signals of daily concentrations into the components listed above. The methods are described in detail in the supporting information; however, key differences are described here. The models take the forms:

$$ln[C_{MKZ}(t)] = \hat{C}^{LT}_{MKZ}(t) + \hat{C}^{S}_{MKZ}(t) + \hat{C}^{WH}_{MKZ}(t) + \hat{C}^{STM}_{MKZ}(t) + \hat{C}^{WN}_{MKZ}(t) \qquad \text{(Eq. 4.6)}$$

$$ln[C_{KEA}(t)] = \hat{C}^{LT}_{KEA}(t) + \hat{C}^{S}_{KEA}(t) + \hat{C}^{WH}_{KEA}(t) + \hat{C}^{STM}_{KEA}(t) + \hat{C}^{WN}_{KEA}(t) \qquad \text{(Eq. 4.7)}$$

The hats indicate an estimated value. In both methods, a $KZ_{365,3}$ filter is used to remove the LT signal. The averaged output—on a day-of-year basis—from a second filter, $KZ_{15,5}$, represents the S component in the MKZ method. The method is named for the use

of multiple KZ filters to separate the detrending components. For the KEA method, S is calculated by averaging the resulting signal after removing LT by date of year, which takes into account holidays that occur on a different date each year (e.g. Thanksgiving and Memorial Day). A result of this averaging is that the weekly signal, WH, is combined with the S signal in the KEA method.

LT and S are subtracted from the observations in the MKZ method, and LT, S, and WH are subtracted from the observations in the KEA method. The remaining signals (WH, STM and WN in the MKZ method and STM and WN in the KEA method) are estimated in regressions. The covariates included in the MKZ regression are (see Table 4-2 for variable definitions): ΔT^m, $(\Delta T^m)^2$, $(\Delta T^m)^3$, ΔWS, ΔRH, ΔRF, $T^M * \Delta T^m$, $T^M * \Delta RH$, ΔT^m_{-1}, ΔWS_{-1}, ΔRH_{-1}, RF_{-1}, ΔT^m_{-2}, ΔWS_{-2}, ΔRH_{-2}, RF_{-2}, weekday indicators, weekday indicators times maximum temperature, month indicators, and holiday indicators. Subscripts represent number of lagged days. In the KEA regression, only the meteorology covariates are used (i.e. none of the indicator variables). Values are provided for the full set of variables are described in the supporting information, and for the regression coefficients in Tables S1 and S2 in the Supplemental Material. The regressions are different in that weekday and holiday effects are modeled explicitly in the MKZ method, while they are not in the KEA method, and $\Delta M(t)$ are estimated slightly differently (see supplementary material).

4.5 Results

4.5.1 Method evaluation and comparison

Correlations (Table 4-3) show that the MKZ model captures as much or more of the observed variability than the KEA method for all but one of the pollutants (SO_4^{2-}). Most pollutants have slopes and R^2 values of greater than 0.40 across all models. SO_2 is the only exception—neither of the models produce an R^2 greater than 0.30 for SO_2. The concentration of SO_2 that is measured at JST is highly dependent on an electricity generating unit (EGU) plume passing over the monitoring site. This is especially true since most mobile emissions of SO_2 decreased dramatically after rules came into effect that reduced the allowable sulfur content in gasoline and diesel beginning in 1999. This has in turn reduced the variability in the SO_2 signal over time, leading to a problem with fit in both methods.

Table 4-3. Slopes, correlation coefficients (R^2 values), and root mean square deviations (RMSD) for the fit of the KEA and MKZ methods.

	Slope [conc conc^{-1}]		R^2		RMSD [conc]	
	KEA	MKZ	KEA	MKZ	KEA	MKZ
$MD8hO_3$	0.72	0.71	0.62	0.66	11.3	10.7
O_3^M	0.72	0.72	0.67	0.71	11.7	11.0
NO_X	0.52	0.53	0.58	0.64	9.60	9.11
NO_X^{morn}	0.33	0.37	0.44	0.50	36.1	34.4
CO	0.44	0.44	0.53	0.53	73.9	73.9
CO^{morn}	0.28	0.29	0.44	0.46	274	271
SO_2	0.28	0.21	0.19	0.23	13.2	12.9
$PM_{2.5}$	0.46	0.47	0.49	0.49	4.85	4.82
SO_4^{2-}	0.41	0.38	0.50	0.48	1.92	1.97
EC	0.38	0.38	0.43	0.45	1.37	1.36
OC	0.38	0.39	0.43	0.48	0.57	0.56

A holdout analysis is a type of cross-validation used to investigate the robustness of a model to missing data. 30 holdout tests were performed by training the model on the data with 10% of the observations removed. The sampling was performed independently for each holdout test, meaning it is possible data points were held out for multiple tests. For each test, the remaining 90% of the data was used to estimate each component of the detrending (i.e. LT, S, WH, STM, and WN). The LT and S components are estimated using the KZ filter, which can handle missing data, and still assigns an estimate to missing days. In the KEA method, the WH component is estimated along with S, and STM is estimated using the best fit of the linear model. In the MKZ method, the WH and STM components are both estimated using fits of a linear regression. The linear regressions are fit using the 90% remaining observations, and regression parameters are used to estimate the withheld 10%. The predicted values are compared with observed using the root mean square deviation (Table S2 in Supplementary Information).

The average root mean square deviation (RMSD) was lower for 7 of the 11 pollutant metrics for the MKZ method, and for only one (OC) in the KEA method (p-value < 0.05). Differences in the rest of the pollutants were not statistically significant. The RMSD for each species describes the WN component of the detrending, so a smaller RMSD means more of the variability is captured by the model.

Monitor design values are used to assess attainment under the National Ambient Air Quality Standards (NAAQS). For ozone, the design value is defined as the three-year

average of the annual fourth highest MDA8h ozone. For $PM_{2.5}$, the design value is defined as the three-year average of the fourth highest 24hr average $PM_{2.5}$. The observed design values for O_3 plotted in Figure 4-2 show a decreasing trend. After detrending with both the KEA and MKZ methods, however, design values show a greater decreasing slope. Results from the two models are similar, and provide evidence that meteorological fluctuations contributed to higher design values in most of the years in the current study. While meteorology can both increase and suppress ozone levels, most of the highest days (which contribute to the design value) are on hot summer days with meteorological conditions that are conducive to ozone formation. When this effect is removed in the detrending, the design values decrease (Figure 4-2).

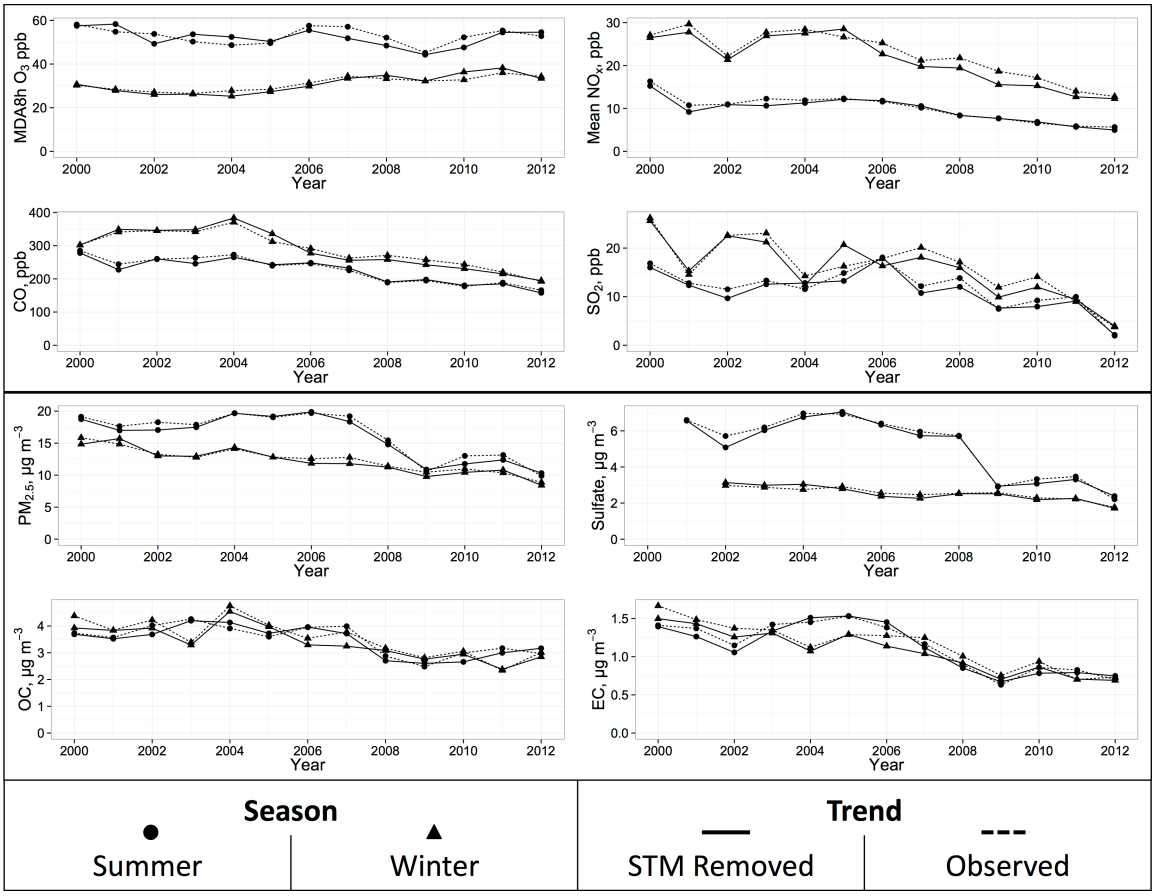


Figure 4-1. The raw and meteorologically-detrended summer and winter means. Winter averages are taken from January 1st to March 31st, and Summer averages are taken from July 1st to September 30th. Gaseous species are above and PM species are below.

Design values for detrended $PM_{2.5}$ from 2003 to 2007 are less than the observed. For the other years, both detrending methods are similar to the observed.

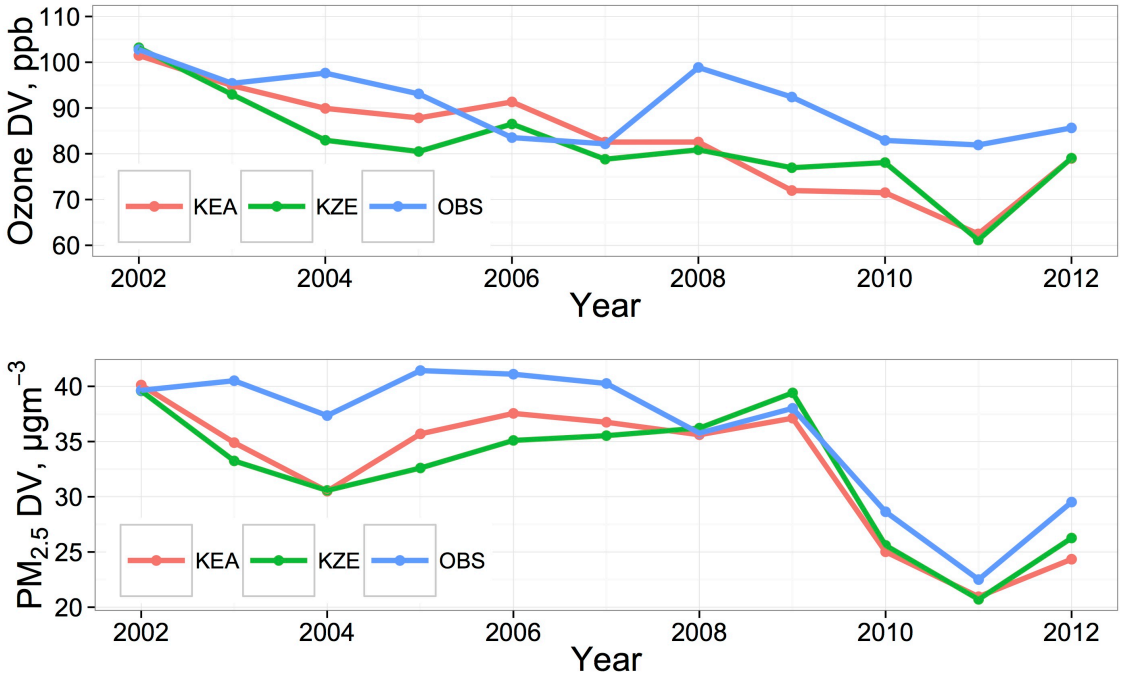


Figure 4-2. Monitor design values for O_3 and $PM_{2.5}$. For O_3, the design value is defined as the three-year average of the annual fourth highest MDA8h O_3. For $PM_{2.5}$, the design value is defined as the three-year average of the fourth highest 24-hr average $PM_{2.5}$.

Results from both models show that average annual ozone concentrations did not change much from 2000 to 2012 in Atlanta (Figure 4-3), although design values of both ozone and $PM_{2.5}$ have decreased substantially (Figure 4-2). Results of the models are similar based on the R^2 values (Table 4-3) and the ability to calculate daily adjustments. Because of the higher R^2 values and results of the holdout analysis, the MKZ method was identified as the most appropriate model. Therefore, a majority of the following discussion addresses results from this model.

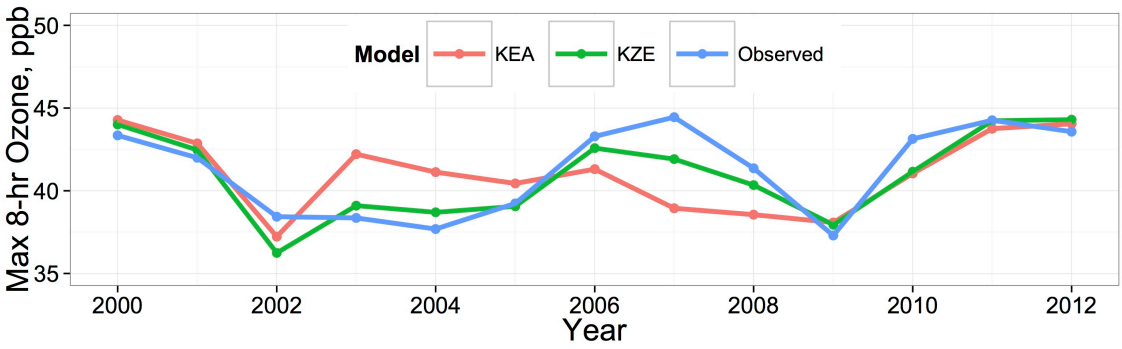


Figure 4-3. The annual average adjusted max of 8-h max ozone levels for Atlanta for the KEA and MKZ, methods compared with the observed.

4.5.2 Effect of meteorology on ozone levels

The magnitude of the daily meteorological contribution to ozone levels are specific to different climate regions (Camalier et al., 2007; Ge go et al., 2007; Rao et al., 1995), and, for Atlanta, approximate a normal distribution (Fig. S1a in Supplementary Material). Fig. S1b in Supplementary Material displays the monthly means and standard deviations of meteorological contribution to ozone levels. In general, meteorological contributions increase in the summer and decrease in the winter. However, the width of the distribution also changes, as the summer sees a wider range of effects than the winter. Two exceptions to this trend are May and July, which both average a near-zero effect and have standard deviations that are smaller than April, June, and August. The cause of this may be traced to the negative association of relative humidity with ozone.

The average contribution of short-term meteorology is 0.9 ppb (near zero because positive and negative contributions balance each other out over the length of the 13-year time series), but the model finds contributions as low as 57.2 ppb and as high as 44.0 ppb (Table 4-4). The mean absolute contribution of daily short- term meteorology is 8.4 ppb. Days with high and low relative humidity show opposite average effects on ozone levels. Though the distributions of high and low relative humidity days have different means, the standard deviations are similar. Therefore, changes in temperature yield a widening of the response distribution, and changes in relative humidity yield a change in the mean direction of the response distribution.

Table 4-4. The mean (μ), standard deviation (σ's), minimum, and maximum of the meteorological contribution of different subsets of the JST ozone data. "High" and "Low" splits are made at the mean value for each variable. MAC is the mean absolute contribution of meteorology, or the average absolute STM for each data transect. The differences in the means are statistically significant (as determined by a Student's t-test) for temperature and relative humidity subsection. Units are ppb.

	μ	σ	Min	Max	MAC
All	0.9	10.3	-57.3	44.0	8.4
Temperature					
High	1.5	12.3	-57.3	44.0	9.8
Low	0.2	8.7	-43.2	28.6	6.5
Relative humidity					
High	-2.0	9.7	-44.9	35.8	7.5
Low	3.2	11.2	-57.4	44.0	9.1

4.5.3 Meteorologically detrended species metrics

Meteorologically detrended and raw summer and winter means have decreased in Atlanta for all pollutants in this analysis over the study period (Figure 4-1). The MKZ detrending process has the effect of reducing the means and variances in the observations of all pollutants, though changes in the variances are more substantial than changes in the mean (Fig. S2 in Supplemental Material) This is an expected result due to the added variability that meteorological fluctuations add to observed concentrations of pollutants.

Detrended winter averages for ozone increase by almost 20 ppb between 2005 and 2011 (Figure 4-1). Summer values show more variability between years. Decreases in summer averages are seen from 2001 to 2002, 2004 to 2005, and 2006 to 2009. However, these decreases are each followed immediately by one or two years of increases in ozone levels. While the overall trend in ozone between 2000 and 2012 shows a slight decrease, it is not nearly as distinct as in the primary pollutants or $PM_{2.5}$.

Both SO_2 and SO_4^{2-} show large decreases in winter concentrations between 2000 and 2001, then remain stagnant until 2008. Concentrations then decrease steadily until 2012, with an exception in 2010. Overall, changes in SO_4^{2-} concentrations mirror changes in SO_2 concentrations with a slightly dampened response. Winter $PM_{2.5}$ shows a near-steady decrease since 2001. Summer values, on the other hand, show little change before 2006 followed by a large drop between 2007 and 2009. Average values in the summer of 2012 were 11 μg m^{-3} less than average values in 2000.

4.5.4 Atmospheric response to emissions changes

In order to identify changes in air quality due to emissions changes, fluctuations from both short-term meteorology and seasonal variations must be accounted for. Seasonal variations tend to dominate the variability in concentration signals, and inhibit the ability to observe short-term changes. To this end, the seasonal (S) portion of the detrended concentration signal is removed. The remaining signal is then averaged over different time periods (1, 2, 3, 4, 6, and 12 months). For each averaging time, adjacent mean concentrations are compared with a Student's t-test, using the Bonferroni method to correct each p-value. This correction accounts for the increased error associated with performing a large number of significance tests (Hastie et al., 2009). All averaging periods are shown for SO_2 (Figure 4-4), along with annual averages for the remaining pollutants (Figure 4-5). See Supplementary material (Figs. S3-S6) for plots of all averaging periods for NO_X, SO_4^{2-}, $PM_{2.5}$, and O_3.

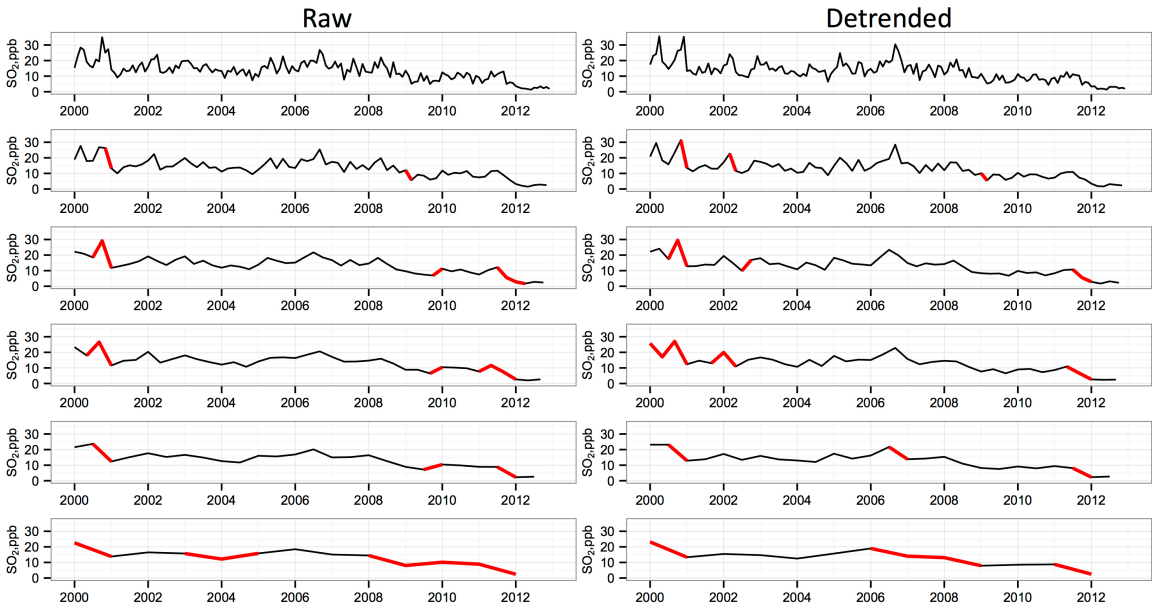


Figure 4-4. Averaged SO_2 concentrations at different averaging periods. Raw concentrations are on the left, and meteorologically detrended concentrations are on the right. The averaging times are (top to bottom): 1, 2, 3, 4, 6, and 12 months. Bold red lines indicate a statistically significant ($p < 0.05$) change in the average concentration.

Meteorological detrending increases the ability to observe a statistically significant change in the observed NO_X, SO_2, and SO_4^{2-} concentrations at shorter averaging periods. Results of the signal processing for SO_2 (Figure 4-4) show that averaging times of 1 and 2

months are still too short to observe meaningful statistically significant (i.e. p-value < 0.05) changes in concentration. Statistically significant peaks in the 1- and 2- month averaging times may be the product of spikes caused by local disturbances. At averaging times of 3 months and greater, however, differences in the meteorologically detrended concentrations become statistically significant and relatable to known emissions changes. At 3 months and higher, both the detrended and raw signals show significant decreases from 2000 to 2001 and 2011 to 2012. Raw concentrations show a significant increase in SO_2 between 2009 and 2010 for averaging periods of 3 months and greater. After detrending, this increase only appears in the 3-month average. At six months averaging periods and longer, no increases are significant.

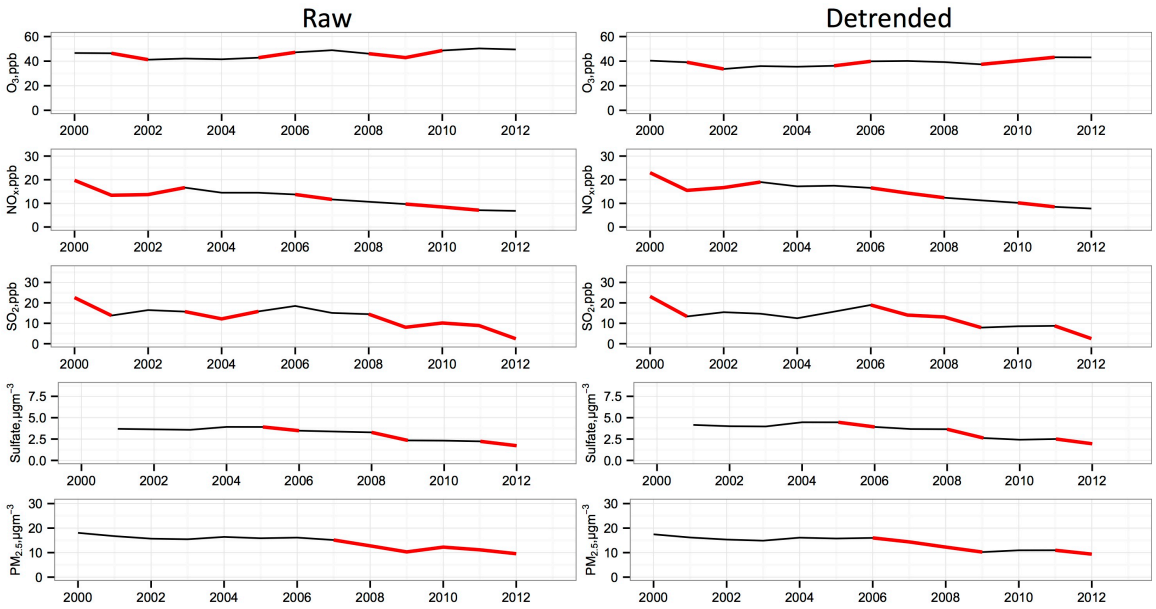


Figure 4-5. Averaged annual concentrations of 5 pollutants. Raw concentrations are on the left, and meteorologically detrended concentrations are on the right. Bold red lines indicate a statistically significant (p < 0.05) change in the annual average concentration.

Detrending of NO_X concentrations (Fig. S3 in Supplementary Material) also shows differences from the measured concentrations. Again, averaging times of 2 months and less capture too much of the noise to show statistically significant differences. First, changes in 2001 and 2007 are significant for averaging times as low as 3 months. These changes are only significant in averaging times of 6 months or greater when meteorological fluctuations are not accounted for. Annual averages show significant changes in the detrended signal in 2007 and 2008 that are not present in the raw signal. Meteorological detrending reduces

the standard deviation of the daily NO_X signal (over the 13 year time period) from 15.8 ppb to 12.8 ppb.

$PM_{2.5}$ and ozone are inherently more difficult to analyze using these signal processing techniques. Part of this difficulty arises from the fact that a large portion of $PM_{2.5}$ and all ozone are secondary pollutants whose primary constituents are emitted from a number of sources. When the seasonal signal is removed from the $PM_{2.5}$ and ozone signals, many statistically significant changes exist that are not easily linked to known emissions changes (Figs. S5 and S6 in Supplementary Material). Annual averages for $PM_{2.5}$ and ozone are exceptions. Statistically significant changes in $PM_{2.5}$ occurred in years from 2007 to 2009 and 2011 to 2012. These years align with significant decreases in sulfate concentrations, which makes sense since sulfate is a major constituent of $PM_{2.5}$ in the southeast. Removing the seasonal signal from the ozone concentrations yields a signal that changes phase (i.e. the signal peaks in the summers at the beginning of the data record but changes to the winter at the end) and in the middle of the data record (Fig. S6 in Supplementary Material). This is due to the narrowing distribution of ozone levels (Figure 4-1). The result of this is that averaging periods below 12 months have little meaning, since the change in phase drives the majority of the change in average concentrations at smaller periods.

4.6 Discussion

4.6.1 Other methods considered

Initially, attempts were made to use band-pass filtering in frequency space (by employing a Fast Fourier Transform (FFT)) to isolate fluctuations at each of the periods in (Eq. 4.2). Results of the FFT algorithm, however, are highly dependent on the length of the dataset, which reduces the ability to generalize the method to other locations with more or less data. Further, the algorithm does not readily handle missing data, which necessitates filling the missing days with averages or linearly interpolating between days. Given the special treatment required to perform FFT filtering, this method was viewed as less attractive, similar to the results reported in Eskridge et al. (1997).

A fourth method that was taken into consideration in the analysis is reported in Camalier et al. (2007). This method was developed to meteorologically detrend 8-h ozone estimates, and results for cities across the United States are reported at http://

www.epa.gov/airtrends/weather.html. However, because the model was developed with the intent of calculating annual adjustments to summer averages, and not day-to-day adjustments, the method was assessed, but is not used further here. Results reported on the EPA website are compared with results from the MKZ and KEA methods.

The MKZ and KEA methods produce detrended annual summer means (Fig. S7 in Supplementary Material) that are comparable with those reported by the EPA at http://www.epa.gov/airtrends/ weather2012/atlaGA.jpeg, accessed 27 May, 2015. Actual values were not available, so the EPA values were estimated via visual inspection. The MKZ and KEA methods estimate an average meteorological impact on summer mean ozone of the same sign in 9 out of 13 years of this study. In three of the years that are not adjusted in the same direction (2000, 2003, and 2009), all three methods estimate small adjustments (2-5 ppb), so the differences may be due to outlier days or variations attributable to different monitor locations.

4.6.2 Comparison with emissions estimates

Reasons for the reductions in primary pollutants can be attributed to a number of factors, including economic activity, fuel costs, and regulatory programs. A number of local and national programs were instituted during this time period. These programs include those targeting point sources (e.g. Phase 2 of the Acid Rain Program; the NO_X Budget Trading Program, Georgia was not included in this program, but neighboring states were; the Clean Air Interstate Rule, and the Georgia Multipollutant Control Rule for Electricity Utility Steam Generating Units) and mobile sources (e.g. Tier 2 Vehicle and Gasoline Sulfur Program, the 2007 Heavy-Duty Highway Rule, the Mobile Source Air Toxics Rule, and the Georgia Gasoline Marketing Rule) (Georgia Department of Natural Resources Environmental Protection Division Air Protection Branch (2013); Hubbell et al., 2009; Morgenstern et al., 2012). Pachon et al (2012) showed a linear relationship between estimated mobile emissions levels from 1999 to 2007 and measured CO, NO_X, and elemental carbon concentrations (Pachon et al., 2012).

Seasonal variability in NO_X emissions between summers and winters from 2003 to 2008 provide an interesting point of comparison to detrended concentrations. These seasonal emissions controls were required by Georgia in an effort to reduce summer ozone levels. Starting in 2009, the state required year-round controls on NO_X. A comparison of

Fig. S8 in Supplemental Material with NO_X concentrations in Figure 4-1 shows that EGU summer controls for NO_x had little effect on the detrended summer NO_X concentrations, which increase each year between 2003 and 2005. Decreases in detrended summer and winter NO_X concentrations are observed after 2005. This provides evidence that emissions from local EGUs have a smaller influence on NO_X concentrations in downtown Atlanta than other sources.

The detrended ozone levels, apparently in contrast to its reduced precursor concentrations, do not decrease nearly as much (4% or 2.2 ± 2 ppb) in the summer between 2000 and 2012. Winter averages increased over this time period by 12% (3.8 ± 2 ppb). These results corroborate recent findings in Blanchard et al. (2010) that the ozone response in Atlanta to NO_X levels is less in magnitude than other factors, such as meteorology and ambient levels of non-methane organic carbons (NMOCs).

Many of the regulatory policies established during this period have targeted the highest ozone days. In particular, from 2003 to 2008, seasonal NO_X controls significantly reduced EGU NO_X emissions during the summer ozone season (May-September, Fig. S8 in Supplemental Material). From 2009 on, these controls were kept in place all year. This change in seasonal emissions is used to assess if detrended ozone concentrations can be used to evaluate the impact of these controls on high ozone levels. Days with detrended concentrations greater than 60 ppb were analyzed separately in an attempt to assess how well the impact of those decreases can be identified. Annual averages of concentrations on these days exhibited a decreasing trend over the time period on average, however, inter-annual variability made it difficult to identify if the seasonal EGU emissions had an appreciable impact. Longer-term (3+ year) averages are more stable and do show a marked decrease from the period without seasonal controls to the first four years after seasonal controls were put in place (Table 4-5). The first averaging period (2000-2002) is before widespread use of seasonal NO_X controls on EGUs in the region, the second (2003-2005) and third (2006-2008) are when seasonal controls were instituted, and the fourth (2009-2012) is when the controls were employed year-round. Average high-ozone day concentrations decrease across the first three periods, and increase slightly in the most recent period. The number of days that saw ozone concentrations greater than 60 ppb decreased significantly over the final three periods. 60 ppb was chosen as a cutoff because

of its relevance to current discussions of ozone standards in the United States. The results are reflected in distributions of all ozone levels (Figure 4-6); the distribution narrows as new controls are implemented. Days of very high or very low ozone are increasingly rare as controls are installed across a number of sectors.

Table 4-5. Averages concentration of O_3 on days it exceeded 60 ppb and average number of days per year in 4 different periods.

Period	Average [ppb]	# days
2000-2002	72.4	96
2003-2005	70.2	110
2006-2008	66.9	57
2009-2012	68.6	45

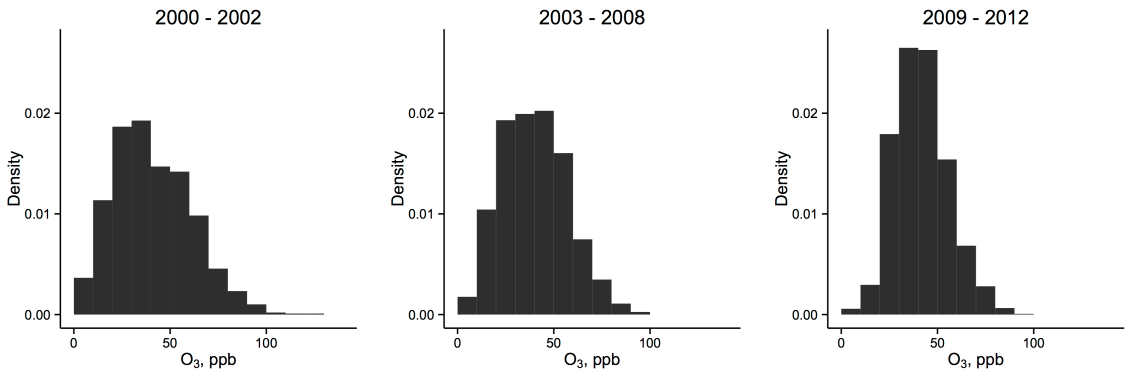


Figure 4-6. Normalized densities of concentrations during 3 subsets of the study period.

4.6.3 Atmospheric response to emissions changes

Because of the range of time scales that are important in atmospheric dynamics, it can be difficult to quantify the effect of a specific control strategy on ambient concentrations. A number of factors regarding the nature of how controls are put in place can also affect the ability to attribute changes in concentration levels at a central monitor to emissions changes. First, regulatory programs are implemented in phases, e.g. Phase I (1995) and Phase II (2000) of the Acid Rain Program. Second, more controls are generally required on new emitters (e.g. automobiles, EGUs, gas stations, etc.), and the effects are not fully realized until the fleet turns over. Third, multiple regulatory programs are implemented across many domains simultaneously, making it difficult to determine which control is attributed to the change in ambient concentration. Fourth, local sources

contribute disproportionally more to the variability in a measured concentration at a central monitor (e.g. JST). The time scales analysis shown in the results and discussed further here seeks to attribute changes in ambient concentrations to actions that have reduced emissions.

Annual mean detrended SO_2 concentrations show three periods of significant decreases (2000-2001, 2006-2009, and 2011-2012). In the 6 month averaging, the change between 2006 and 2007 is significant. This is the same year that total sales of No. 2 diesel fuel sold by the prime supplier in Georgia changed from 91% Low Sulfur diesel (500 ppm S) and 0% Ultralow Sulfur diesel (15 ppm S) to 35% and 58%. The remaining diesel sold was High Sulfur diesel (5000 ppm S). By 2012, 100% of the diesel sold in Georgia was Ultralow Sulfur diesel (U.S Energy Information Administration, 2014). EGU SO_2 emissions fell between 2008 and 2009. This also con- tributes to the significant differences in ambient concentrations observed between these years (Figure 4-5). The change between 2011 and 2012 is statistically significant for both detrended and raw concentrations for an averaging period of 3 months and greater. Part of this decline can be attributed to the conversion of plant McDonough (a power plant located 10 km northwest of JST) to natural gas between these two years.

One of the significant changes in the NO_X record (Fig. S3 in Supplementary Material) is the change between 2000 and 2001, which is detected at an averaging time of 3 months and greater. NO_X emissions from Plant McDonough were reduced by 20% between 2000 and 2001. Over the total 20-county Atlanta $PM_{2.5}$ nonattainment area, EGU NO_X emissions were reduced by 9.6%. These changes, combined with emissions reductions in other sectors (such as mobile), contributed to a change in both annual average raw and detrended NO_X concentrations between 2000 and 2001 of 32%.

Annual detrended averages of sulfate display significant changes in the same years as SO_2 (Figure 4-5). This is expected because sulfate is formed in the atmosphere from SO_2. However, since sulfate must undergo chemical transformations, it is considered a regional pollutant, meaning local sources are not expected to have as much of an impact on sulfate than on SO_2. The fact that the two align provides some evidence that strategies to reduce SO_2 emissions in Atlanta influence sulfate concentration, but is also a

consequence of the fact that similar regulations have been implemented across the entire region.

The lowest annual average detrended ozone level was observed in 2002, and the two highest values were in 2011 and 2012 (Figure 4-5). Statistically significant increases between 2009 and 2011 are present even as NO_X concentrations fell. Annual average ozone levels are increasing because winter concentrations are rising faster than summer levels are falling (Figure 4-1). The dramatic rise in winter concentrations provides evidence that winter ozone levels are NO_X inhibited, i.e. decreasing NO_X emissions yield higher ozone concentrations. High levels of ozone in the summer show the opposite effect.

4.7 Summary/Conclusions

Meteorological detrending is applied to remove the effect of meteorological fluctuations on measured gaseous and particulate matter pollutants. The daily impact can be substantial, with a mean absolute contribution of 8.4 ppb, with range of 57.2 to 44.0 ppb for ozone. The KZE method builds on previous methods (including the KEA method) to quantify the impact of short-term meteorology on air pollution concentrations. Both methods yield similar results to each other when assessing design values at JST for ozone and $PM_{2.5}$. This investigation shows that the new method is able to do as well as or better than the KEA method for all pollutants except SO_4^{2-} in the model fit, and OC in the holdout analysis (Figure 4-4). Detrended summertime ozone averages align well with previous results reported by the EPA using the method detailed in Camalier et al. (2007).

In NO_X, SO_2, $PM_{2.5}$, and SO_4^{2-} time series with the seasonal variation removed, statistically significant changes in annual averaged detrended concentrations coincide well with changes in emissions from EGU's and mobile sources. Annual average detrended SO_2 shows statistically significant changes when sulfur concentration in gasoline and diesel was reduced from 2006 to 2008, and when the fuel used at a nearby EGU (Plant McDonough) was changed from coal to natural gas in 2012. Attempts were made to parse out statistically significant changes in NO_X concentrations that reflected summer NO_X controls on Atlanta-area EGU's from 2003 to 2009, but these could not be observed. This is due, in part, to the smaller impact local EGU emissions have on measured concentrations in the city center compared to the impact of mobile NO_X sources.

While seasonal averaging (Figure 4-4 and Figs. S3 and S4 in Supplementary Material) allows for the identification of many of the significant changes in atmospheric concentrations of NO_X, SO_2, and sulfate, these averaging times are not as useful for ozone and $PM_{2.5}$. Further, in contrast to the primary pollutants and $PM_{2.5}$, statistically significant changes in ozone concentrations did not coincide with any of the emissions changes mentioned above. Ozone is the only pollutant that does not show a major decrease in mean levels over the 13-year study period. Detrended average summer concentrations decreased 4% (2.2 ± 2 ppb) while winter concentrations increased by 12% (3.8 ± 1.4 ppb) between 2000 and 2012. These results suggest that summer peak ozone levels are NO_X limited, while lower ozone levels are radical-limited (NO_X inhibited). The detrending method developed can be used to adjust for the meteorological impacts on a day-by-day basis for use in acute health studies. The MKZ method provides the benefit of daily detrending that also produces annual averages that agree with previously published detrending methods.

4.8 Acknowledgments

This material is based upon work supported by Health Effects Institute and the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1148903.

This publication was made possible, in part, by US EPA grant R834799. This publications contents are solely the responsibility of the grantee and do not necessarily represent the official views of the US EPA. Further, US EPA does not endorse the purchase of any commercial products or services mentioned in the publication.

4.9 References

Atmospheric Research and Analysis, 2014. http://www.atmospheric-research.com/ studies/SEARCH/index.html.

Blanchard, C., Hidy, G., Tanenbaum, S., 2010. NMOC, ozone, and organic aerosol in the southeastern United States, 19992007: 2. Ozone trends and sensitivity to NMOC emissions in Atlanta, Georgia. Atmos. Environ. 44 (38), 4840e4849. http://dx.doi.org/10.1016/j.atmosenv.2010.07.030. http://linkinghub.elsevier. com/retrieve/pii/S1352231010005996.

Brönnimann, S., Neu, U., 1997. Weekend-weekday differences of near-surface ozone concentrations in Switzerland for different meteorological conditions. Atmos. Environ. 31 (8), 1127e1135. http://www.sciencedirect.com/science/article/pii/ S1352231096003111.

Camalier, L., Cox, W., Dolwick, P., 2007. The effects of meteorology on ozone in urban areas and their use in assessing ozone trends. Atmos. Environ. 41 (33), 7127e7137. http://dx.doi.org/10.1016/j.atmosenv.2007.04.061. http://linkinghub.elsevier.com/retrieve/pii/S1352231007004165.

Close, B., Zurbenko, I., 2011. kza: Kolmogorov-Zurbenko Adaptive Filters. http:// CRAN.R-project.org/package1⁄4kza, R package version 2.01.

Cox, W.M., Chu, S.-H., 1993. Meteorologically adjusted ozone trends in urban areas: a probabilistic approach. Atmos. Environ. B. Urban Atmos. 27 (4), 425e434. http://www.sciencedirect.com/science/article/pii/0957127293900193.

Eskridge, R.E., Ku, J.Y., Rao, S.T., Porter, P.S., Zurbenko, I.G., 1997. Separating different scales of motion in time series of meteorological variables. Bull. Am. Meteorol. Soc. 78 (7), 1473e1483. http://dx.doi.org/10.1175/1520-0477(1997)078¡1473: SDSOMI¿2.0.CO;2. http://journals.ametsoc.org/doi/abs/10.1175/1520-0477% 281997%29078%3C1473%3ASDSOMI%3E2.0.CO%3B2.

Flaum, J.B., Rao, S.T., Zurbenko, I.G., 1996. Moderating the influence of meteoro- logical conditions on ambient ozone concentrations. J. Air Waste Manag. Assoc. 46 (1), 35e46. http://dx.doi.org/10.1080/10473289.1996.10467439. http://www. tandfonline.com/doi/abs/10.1080/10473289.1996.10467439.

Gardner, M., Dorling, S., 2000. Meteorologically adjusted trends in UK daily maximum surface ozone concentrations. Atmos. Environ. 34 (2), 171e176. http://dx.doi.org/10.1016/S1352-2310(99)00315-5. http://linkinghub.elsevier. com/retrieve/pii/S1352231099003155.

Ge go, E., Porter, P.S., Gilliland, A., Rao, S.T., 2007. Observation-based assessment of the impact of nitrogen oxides emissions reductions on ozone air quality over the Eastern United States. J. Appl. Meteorol. Climatol. 46 (7), 994e1008. http:// dx.doi.org/10.1175/JAM2523.1. http://journals.ametsoc.org/doi/abs/10.1175/ JAM2523.1.

Georgia Department of Natural Resources Environmental Protection Division Air Protection Branch, 2013. Rules for Air Quality Control.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning. Springer, New York.

Hogrefe, C., Rao, S.T., 2000. Interpreting the information in ozone observations and model predictions relevant to regulatory policies in the eastern United States. Bull. Am. Meteorol. Soc. 2083e2106. http://cat.inist.fr/?aModele1⁄4afficheN& cpsidt1⁄41481524.

Hubbell, B.J., Crume, R.V., Evarts, D.M., Cohen, J.M., 2009. Policy monitor: regulation and progress under the 1990 clean air act amendments. Rev. Environ. Econ. Policy 4 (1), 122e138. http://dx.doi.org/10.1093/reep/rep019. http://reep. oxfordjournals.org/cgi/doi/10.1093/reep/rep019.

Kuebler, J., van den Bergh, H., Russell, A.G., 2001. Long-term trends of primary and secondary pollutant concentrations in Switzerland and their response to emission controls and economic changes. Atmos. Environ. 35 (8), 1351e1363. http://dx.doi.org/10.1016/S1352-2310(00)00401-5. http://linkinghub.elsevier. com/retrieve/pii/S1352231000004015.

Morgenstern, R.D., Harrington, W., Shih, J.-S., Bell, M.L., 2012. Accountability anal- ysis of title IV phase 2 of the 1990 clean air act amendments. Res. Rep. (Health Effects Inst.) 168.

Pachon, J., Balachandran, S., Hu, Y., Mulholland, J.A., Darrow, L.A., Sarnat, J.A., Tolbert, P.E., Russell, A.G., 2012. Development of outcome-based, multipollutant mobile source indicators. J. Air Waste Manag. Assoc. 62 (11). http://www.tandfonline.com/doi/abs/10.1080/10473289.2012.656218.

R Core Team, 2012. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. http://www.R-project.org/.

Rao, S., Zalewsky, E., Zurbenko, I., 1995. Determining temporal and spatial variations in ozone air quality. J. Air Waste Manag. Assoc. 45 (1), 57e61. http://www.tandfonline.com/doi/abs/10.1080/10473289.1995.10467342.

Rao, S., Zurbenko, I., 1994. Detecting and tracking changes in ozone air quality. J. Air Waste Manag. Assoc. 44, 1089e1092. http://www.tandfonline.com/doi/abs/10.1080/10473289.1994.10467303.

Rao, S., Zurbenko, I., Neagu, R., 1997. Space and time scales in ambient ozone data. Bull. Am. Meteorol. Soc. http://cat.inist.fr/?aModele1⁄4afficheN&cpsidt1⁄42857639. U.S Energy Information Administration, 2014. Prime Supplier Sales Volumes. http://www.eia.gov/dnav/pet/pet_cons_prim_dcu_SGA_a.htm.

USEPA, 1999. Impact Analysis Control of Air Pollution from New Motor Vehicles: Tier 2 Motor Vehicle Emissions Standards and Gasoline Sulfur Control Re- quirements. Tech. rep.. United States Environmental Protection Agency https://www.gpo.gov/fdsys/pkg/FR-2000-02-10/pdf/00-19.pdf

USEPA, 2000. Regulatory Impact Analysis: Heavy-duty Engine and Vehicle Stan- dards and Highway Diesel Fuel Sulfur Control Requirements. Chapter 3. Tech. rep.. United States EPA.

USEPA, 2009. NOx Budget Trading Program e Basic Information. Tech. rep., pp. 1e8 http://www.epa.gov/airmarkets/progsregs/nox/docs/NBPbasicinfo.pdf.

van Erp, A.M.M., O'Keefe, R., Cohen, A.J., Warren, J., 2008. Evaluating the effec- tiveness of air quality interventions. J. Toxicol. Environ. Health Part A 71 (9e10), 583e587. http://dx.doi.org/10.1080/15287390801997708. http://www.ncbi.nlm.nih.gov/pubmed/18569630.

Vijayaraghavan, K., DenBleyker, A., Ma, L., Lindhjem, C., Yarwood, G., 2014. Trends in on-road vehicle emissions and ambient air quality in Atlanta, Georgia, USA, from the late 1990s through 2009. J. Air Waste Manag. Assoc. 64 (7), 808e816. http://dx.doi.org/10.1080/10962247.2014.892039. http://www.tandfonline.com/doi/abs/10.1080/10962247.2014.892039.

Zurbenko, I., 1991. Spectral analysis of nonstationary time series. Int. Stat. Rev. 39, 163e173.

CHAPTER 5. ASSESSING REGULATORY IMPACTS ON OZONE AND $PM_{2.5}$ USING EMPIRICALLY- AND AIR QUALITY MODEL-DERIVED POLLUTANT SENSITIVITIES[4]

5.1 Abstract

Since the 1990 Clean Air Act Amendments, the USA has seen dramatic decreases in air pollutant emissions from a wide variety of source sectors, which have led to changes in pollutant concentrations: both up and down. Multiple stakeholders, including policy-makers, industry, and public health professionals, seek to quantify the benefits of regulations on air pollution and public health, a major focus of air pollution accountability research. Two methods, one empirical, the other based on a chemical transport model (CTM), are used to calculate the sensitivities of ozone (O_3) and particulate matter with diameters less than 2.5 μm ($PM_{2.5}$) to electricity-generating unit (EGU) and mobile source emissions. Both methods are applied to determine impacts of controls on daily concentrations (which are important in assessing acute health responses to air pollution), accounting for nonlinear, meteorologically, and emission-dependent responses of pollutant concentrations. The statistical method separates contributions of nearby EGU, regional EGU, and mobile source emissions on ambient city-center concentrations. Counterfactual emissions, an estimate of emissions under a scenario where no new controls were implemented on local EGU sources after 1995, regional EGUs after 1997, and mobile sources after 1993, are combined with these sensitivities to estimate counterfactual concentrations that represent what daily air quality in Atlanta, GA would have been had controls not been implemented and other emissions-reducing actions not been taken. Regulatory programs are linked with reduced peak summertime O_3, but have had little effect on annual median concentrations at the city-center monitoring site, and led to

[4] A modified version of this chapter has been published as: Henneman, L.R.F., Chang, H., Lavoué, D., Mulholland, J., Russell, A. (2017). "Assessing regulatory impacts on ozone and $PM_{2.5}$ using empirically- and air quality model-derived pollutant sensitivities." *Air Quality, Atmosphere, and Health.* doi:10.1007/s11869-017-0463-2. The online version of this article contains supplementary material, which is available to authorized users.

increases in pollutant levels under less photochemically-active conditions. The empirical method and the CTM method found similar relationships between ozone concentrations and ozone sensitivity to anthropogenic emissions. Compared to the counterfactual between 2010 and 2013, the number of days on which O_3 ($PM_{2.5}$) concentrations exceeded 60 ppb (12.0 μg m^{-3}) was reduced from 396 to 200 (1391 to 222). In 2013, average daily ambient O_3 and $PM_{2.5}$ concentrations were reduced by 1.0 ppb (2%) and 9.9 μg m-3 (48%), respectively, and fourth highest maximum daily average 8-h O_3 was reduced by 14 ppb. Comparison of model-derived sensitivities to those derived using empirical methods show coherence, but some important differences, such as the O_3 concentration where the sensitivity to NO_X emissions changes sign.

5.2 Introduction

In response to the 1990 Clean Air Act Amendments and other measures, the US Environmental Protection Agency (US EPA), states, and local agencies have implemented a number of policies that are designed to decrease emissions of pollutants linked with adverse health outcomes (e.g., USEPA 1999b, 2009, 2005, 2000a; Georgia EPD 2013). Goals of the various programs include bringing areas into compliance with the National Ambient Air Quality Standards (NAAQS) and reducing air toxics emissions (NAP 2004). Of the six criteria pollutants for which NAAQS are promulgated, ozone (O_3) and fine particulate matter have proven the most difficult to reduce below the standards levels. As of July 2016, 214 and 87 full or partial counties were designated in non-attainment for O_3 and particulate matter ($PM_{2.5}$ or PM_{10}), respectively, compared to 26 and 47 counties for lead and sulfur dioxide (SO_2), and none for carbon monoxide or nitrogen dioxide (NO_2) (http://www3. epa.gov/airquality/greenbook/ancl.html, *accessed 13 July, 2016*). These designations are based on the previous O_3 standard (0.75 ppb, which was changed to 0.70 ppb on 1 October, 2015) since the EPA will not update its non-attainment designations for the new standard until late 2017 (U.S. EPA 2015b).

In assessing the effectiveness of regulatory programs, a number of challenges arise (HEI 2003; van Erp et al. 2008). First, multiple regulations may be implemented at or near the same time across different sectors, making it difficult to disentangle the effects of a specific regulation from others. Second, controls may not have immediate effects, instead yielding an increasing effect over time, e.g., cleaner cars entering the vehicle fleet and

replacing older, more polluting cars. Over those longer periods, other long-term changes in weather, land use, and other source emissions may occur. Third, compliance timelines and effectiveness may not be uniform over all targets of a regulation. In regard to stationary sources, operators may take actions such as retrofitting a plant to a new fuel, installing controls, or switching load between different plants for a variety of reasons besides current regulations, including anticipated future regulations, changing demand, and fuel costs (Georgia Power 2007; van Erp et al. 2008). Further, chemical reactions between certain atmospheric pollutants may result in different impacts on ambient concentrations from the reduction of multiple pollutants (NO_X = NO + NO_2 and VOCs—volatile organic compounds—in the case of ozone) than to reductions of a single pollutant (e.g., Cohan et al. 2005; Seinfeld and Pandis 2006). Further still, variations in meteorological conditions lead to differences in reaction rates, atmospheric transport, and deposition that affect pollutant concentrations.

The present study uses ambient air pollution concentrations, measured emissions from power plants (also called electricity generating units—EGUs), and modeled mobile source emissions in Atlanta, Georgia from 1999 to 2013 to develop counterfactual time series of ozone and $PM_{2.5}$ that assume no additional policies were implemented after 1995 on local EGUs, 1997 on regional EGUs, and 1993 on mobile sources. Important EGU regulatory programs implemented during this period include (year the program began) the Acid Rain Program (1995), the Clean Air Interstate Rule (CAIR—2008), and the Georgia Multipollutant Rule (2009). Mobile source programs include the Georgia Gasoline Marketing Rule (1999), the Tier 2 Vehicle and Gasoline Sulfur Program (2004), and the Heavy-Duty Highway Rule (2007), USEPA (1999a, 2000b, 2005, 2012a; EPD 2014).

Researchers have calculated ambient pollutants sensitivities to emissions using statistical methods (e.g., Blanchard et al. 2010; Harrington et al. 2012) and using first-principles chemical transport models (CTMs), both by brute force (e.g., Digar and Cohan 2010 and Xie et al. 2011) and direct methods (e.g., Dunker 1981, 1984; Cohan et al. 2005; Liao et al. 2008; Hakami et al. 2004). The Community Multiscale Air Quality Model with the Decoupled Direct Method (CMAQ-DDM) and the Comprehensive Air Quality Model with Extensions (CAMx, which includes DDM) are examples of models that can calculate sensitivities directly.

CTMs such as CMAQ-DDM and CAMx offer the benefit of incorporating detailed physics and chemistry parameterizations in the calculations of concentrations and sensitivities. Studies have shown that results are sensitive to uncertainties in meteorological inputs (Appel et al. 2007; Byun et al. 2007; Gilliland et al. 2008), emissions inputs (Byun et al. 2007; Gilliland et al. 2008; Zhang et al. 2015), and the combination of chemical and physical mechanisms employed in the model (Hanna et al. 2001; Appel et al. 2007; Byun et al. 2007; Gilliland et al. 2008). In spite of these uncertainties, CTMs have been demonstrated to accurately simulate ambient concentrations and capture observed trends under changing emissions in dynamic evaluations (Foley et al. 2015b; Foley et al. 2015a; Zhou et al. 2013).

Statistical models do not explicitly utilize information on the detailed physics and chemistry that influence ambient air pollution concentrations. Instead, prior knowledge and empirical information are used to select an initial set of variables that may be associated with the outcome—in this case, ozone and $PM_{2.5}$ concentrations. Physical and chemical processes are captured through empirical relationships between the response variable and model inputs. Care must be taken in the implementation and interpretation of these models, however, as the models require a number of assumptions to be met for the models to be appropriate and their outputs reliable. For instance, collinearity in statistical model inputs can lead to regression coefficients that may not reflect physical reality and may confound the results. If care is taken to account for such issues, statistical models serve as a valuable tool for investigating the relationships between multiple variables. In general, they are less computationally and time-intensive to utilize than CTMs and are driven directly by observations, as opposed to air quality models with uncertain inputs and parameters. A further benefit of not explicitly including physical and chemical parameterizations is that the model may be able to approximate relationships that are not accurately captured in chemical transport models.

Statistical models of ozone have been used to examine how certain meteorological conditions contribute to ozone formation (Bloomfield et al. 1996; Camalier et al. 2007; Henneman et al. 2015; Kuebler et al. 2001), model ozone levels in future climate conditions (Chang et al. 2014), investigate geographic differences in relationships between meteorology and ozone concentrations (Davies and Kelly 1992), and to adjust air quality

model outputs to better match observations (e.g., Hu et al. 2014; Porter et al. 2015). Harrington et al. (2012) used monthly-averaged $PM_{2.5}$ concentrations and power plant emissions in a linear regression model to investigate the effects of regulations imposed under the 1990 Clean Air Act Amendments on $PM_{2.5}$ concentrations in the USA.

This work develops detailed statistical models for assessing daily impacts of controls on both ozone and $PM_{2.5}$, and produces counterfactual time series of pollutant concentrations from 1999-2013. The work is unique in relation to others discussed in the introduction in its use of daily emissions and meteorology in statistical models over such a long period of detailed measurements. Daily results are important both for health analyses and for capturing sub-seasonal responses to emissions controls. We apply and compare results from statistical and CTM-derived approaches.

5.3 Data and Methods

5.3.1 Meteorological and ambient air quality measurements

Air quality and meteorological observation datasets used here are described in detail in Henneman et al. (2015), though for the present work the data time span was expanded by 1 year at the beginning and the end to a range of 1999–2013. Ambient concentrations and meteorological data were obtained from the SouthEastern Aerosol Research and Characterization (SEARCH) network's Jefferson Street (JST) monitoring station (33.777°N, 84.416°W) in Atlanta, Georgia (Hansen et al. 2003). Hourly data were converted to daily metrics using metric-driven averaging times (Table 5-1). Maximum daily 8-h average O_3 (MDA8h O_3) and 24-h average $PM_{2.5}$ are used because they are the standard metrics used for regulatory purposes in the USA, and have been widely used in health impact studies (e.g., Pope et al. 2009; Rich et al. 2012; Garcia et al. 2011). JST is located near downtown Atlanta (Figure 5-1), and represents urban conditions, which may differ slightly from other locations in the broader Atlanta area. In cases of missing data, meteorological observations from JST were supplemented with measurements from Hartsfield-Jackson International Airport (ATL), which is southwest of the city center. Rainfall data came from the airport monitor.

Table 5-1. Daily pollutant species used in detrending analysis (2000-2012). Hourly measurements from JST are converted to daily metrics (# days = 4749).

Metric	Averaging description	Period	# Days	% Missing days
MDA8h O_3	Max of 8h mean	12am-12am	4681	1.43
$PM_{2.5}$	Daily mean	12am-12am	4015	9.41

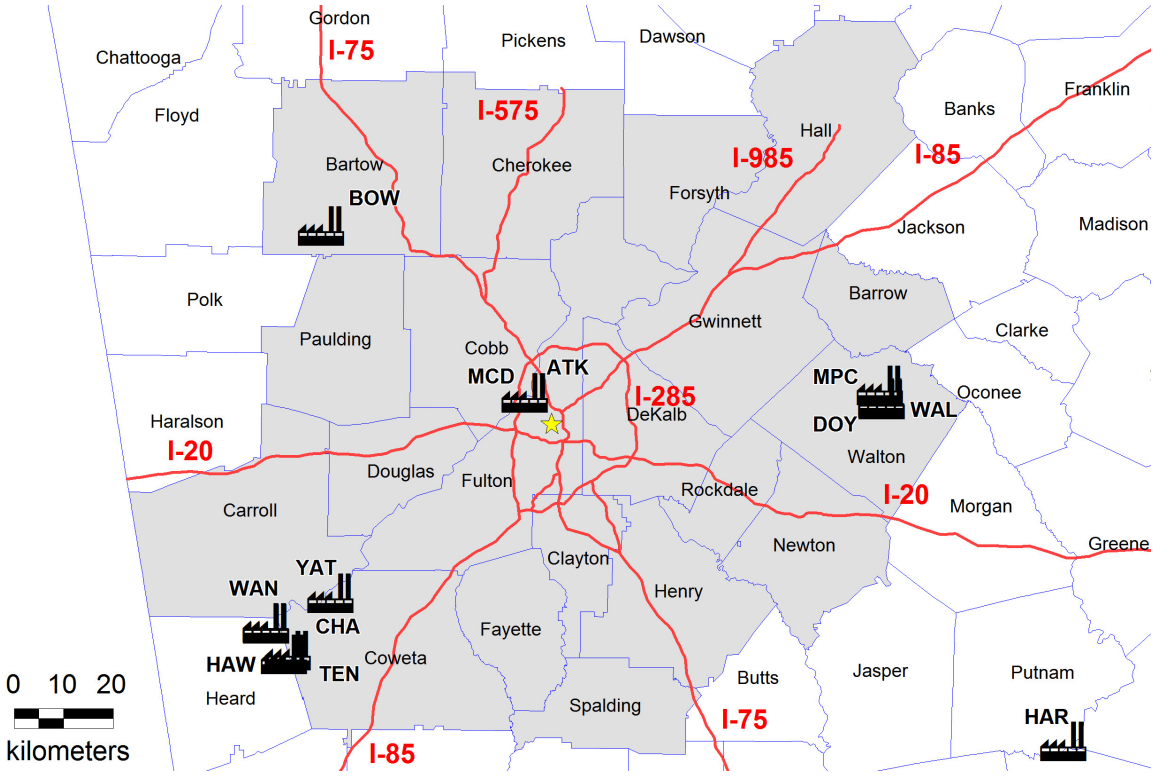


Figure 5-1. The 20-county Atlanta $PM_{2.5}$ Non-Attainment Area (shaded). JST is denoted by the star. Power plants are Atkinson (ATK), Bowen (BOW), Chattahoochee Energy Facility (CHA), Doyle Generating Facility (DOY), Harllee Branch (HAR), Hawk Road Energy Facility (HAW), McDonough (MCD), MPC Generating, LLC (MPC), Tenaska Georgi Generating Station (TEN), Walton County Power, LLC (WAL), Wansley (WAN), and Yates (YAT).

5.3.2 Emissions from mobile sources and EGUs

Mobile source emissions were modeled using the EPA's MOVES2010b software (USEPA 2012c) for the Atlanta 20-county $PM_{2.5}$ Non-Attainment Area (ANAA—Figure 5-1). The ANAA is designated by the US EPA as the area surrounding the city of Atlanta that is in non-attainment of the NAAQS. This area is subject to emissions limits that are stricter than those in the surrounding counties (EPD 2009, 2012). For a discussion of the MOVES model setup and inputs, see the supplemental information.

MOVES estimates mobile emissions using inputs and internal parameterizations that change month-to-month (e.g., temperature and fuel formulation) and year-to-year (e.g.,

vehicle population, inspection and maintenance, and vehicle miles traveled). Consequently, estimated daily emissions often exhibit unrealistic step functions, e.g., between months and years, that should be more gradual. These were corrected using a linear smoothing model that includes linear, squared, and cubed calendar date, weekday/weekend indicators, sine and cosine terms with period of one year, and a time-cosine interaction term (see supplemental information).

Total daily NO_X and SO_2 emissions (tons) and load (MWh) from EGUs in the Southeast region (states included: Alabama, Georgia, Mississippi, North Carolina, South Carolina, and Tennessee) were retrieved from the EPA's Air Markets Program Database (USEPA 2013). Under the Acid Rain Program (ARP) beginning in 1995, EPA has required large emitters to report Continuous Emissions Monitoring (CEM) data. This data was separated into two groups: those within the ANAA (Figure 5-1) and those outside. Daily emissions from all EGUs in each group were summed to represent total EGU emissions. Not all EGUs shown in Figure 5-1 were online during the entirety of the study period. The load for the ANAA plants represents the demand on the suite of plants in Figure 5-1, and not the demand of the greater Atlanta area, which may be met by importing electricity from other counties or across state lines. Further, electricity may be exported from the area to meet demand elsewhere.

Emissions from different locations within the ANAA have different effects on the measured O_3 and $PM_{2.5}$ concentrations at JST, and these effects vary across days due to meteorology (e.g., changing wind direction). Sources inside of the ANAA may have different effects on pollutant concentrations in Atlanta than sources that are farther away (Muller et al. 2009) and including regional emissions separately accounts for this.

5.3.3 Empirical estimates of pollutant sensitivities to source emissions

The empirical sensitivity method developed for the present study has two goals: (1) determine the sensitivity of air pollutant concentrations to emissions changes and (2) use these sensitivities to calculate daily air pollutant concentrations for a counterfactual scenario that assumes no new controls were installed after 1993 in mobile sources and 1995 in EGU sources. Sensitivities were calculated using iteratively weighted least squares regression analysis between observed concentrations and estimated emissions for both ozone and $PM_{2.5}$. Statistical analyses in this work were performed using version 3.2.0 of

the statistical software R; the regressions were fit with the 'glm' command (R Core Team 2015).

5.3.3.1 Empirical ozone sensitivities

Covariates for the empirical ozone model were selected based on results from published literature of pollutant sensitivity analyses (e.g., Cohan et al. 2005; Liao et al. 2008; Seinfeld and Pandis 2006; Xing et al. 2011; Blanchard and Hidy 2005). The original list of covariates included NO_X emissions from EGUs (both within the ANAA and regional, denoted REG), and NO_X and VOC emissions from mobile sources, as well as an interaction between these two. Other covariates included NOx concentration, mean daily wind-speed (WS), temperature (Temp), and relative humidity (RH), and daily rainfall (RF) as a factor (0−1) variable. All four meteorological variables were centered by subtracting their mean to ease the interpretation of model parameters. The sensitivity of the ozone concentrations to emissions is dependent on the level of photochemical activity, which is often characterized by the ozone level in the atmosphere (indeed, the EPA uses the ozone standard as an indicator for atmospheric photochemical oxidants US EPA 2015a). O_3 serves as a proxy for how much OH is available to oxidize a variety of atmospheric constituents, including those that eventually condense to form secondary PM (such as VOCs for secondary organic aerosols and NO_2 for nitrates). Since ozone is the response in the model, however, raw ozone observations cannot be used as a covariate in the model. To account for this, a measure of emissions-independent atmospheric photochemical oxidative state (PS^*) was applied as an effect modifier with multiple emissions covariates.

PS^* was estimated using components from a meteorological detrending method developed to investigate daily impacts of meteorological fluctuations on pollutant concentrations and described in detail in Henneman et al. (2015). In brief, filtering and linear regressions were used to separate different time scales of fluctuations, including long term (period > 1 year), seasonal (period = 1 year), weekly, and short-term meteorological (period < 3 months) contributions. To calculate PS^*, we summed the seasonal (S, which does not vary between years and is synonymous with annual fluctuation) and short-term meteorological (STM) O_3 trends. S was estimated using a Kolmogorov-Zurbenko filter, a low-pass moving average filter, and averaging the output by date-of-year. STM, an estimate of the impact of daily variability in meteorological variables on ozone, was

estimated with a regression of daily fluctuations in multiple metrics: solar radiation (total and daily max) temperature (mean and daily maximum), wind speed (morning and daily means), relative humidity, rainfall, and 1- and 2-day lags of each of these. The sum of S and STM yields a daily metric for the photo-oxidative potential in the atmosphere (Figure 5-2). The resulting metric, PS^*, is higher in the summer than the winter, on warmer days, on drier days, and on days with higher wind speed (likely because surface-level NO, which titrates O_3, is carried away from the city on these days). PS^* was centered by subtracting the mean, so that the average contribution to ozone levels is zero, and was used as an interaction between EGU NO_X, mobile NO_X and VOC, and the interaction between mobile NO_X and VOC emissions.

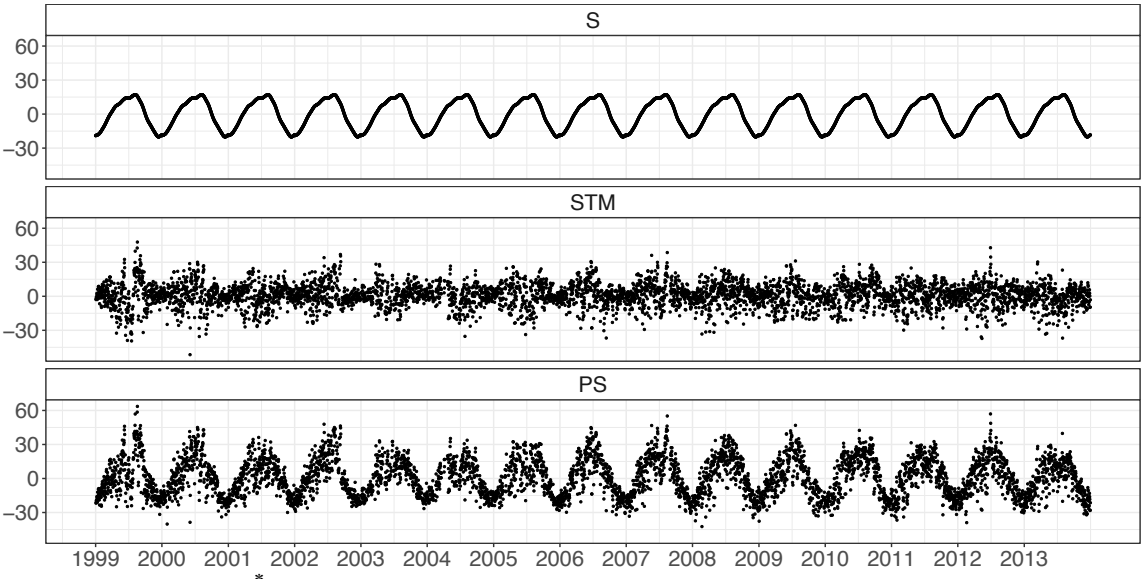


Figure 5-2. PS^* (bottom) and its two components—seasonal fluctuation (S) and short-term meteorological variation (STM) as calculated by Henneman et al. (2015). Units are ppb.

To reduce over-fitting, covariates were removed one at a time from the original list in order of decreasing significance, as measured by the p-values of the parameter error statistics. In general, only regression coefficients significant at the 0.05 level were retained. One exception is the coefficient associated with ANAA EGU NO_X emissions ($p = 0.20$), which was included in the final model because of its known chemical relevance. Modeled mobile emissions exhibit high co-linearity between emitted species (VOC, CO, NO_X) across the time series. Including multiple species in a statistical model, therefore, presents

a problem. Mobile CO emissions contribute to ozone formation (Seinfeld and Pandis 2006), but were excluded from the analysis because of co-linearity with modeled NO_X and VOC emissions, and the chemical action of CO on O_3 formation is similar to that of VOCs. The mobile VOC emissions term was removed because of the co-linearity between modeled mobile NO_X and VOC emissions; however, VOC emissions appear in the model in interaction terms. Model iterations used in the model selection process are provided in the supplemental information. The following is the final model:

$$[O_3] = ß_0 + (ß_1)E_{EGU}^{NO_x} + (ß_2 + ß_3 PS^*)E_{EGU,reg}^{NO_x} + (ß_4 + ß_5 PS^* + ß_6 E_{MOB}^{VOC})E_{MOB}^{VOC} + ß_7 WS + ß_8 Temp + ß_9 RH + ß_{10} RF + \varepsilon_{O_3} \quad \text{(Eq. 5.1)}$$

where $[O_3]$ is MD8hO_3, E_j^k is daily emissions, and ε_{O_3} are the model residuals.

5.3.3.2 Empirical $PM_{2.5}$ sensitivities

The model of $PM_{2.5}$ sensitivities to EGU and mobile emissions is similar to that of ozone. The variable list originally included ANAA and regional EGU NO_X and SO_2 emissions, mobile NO_X, $PM_{2.5}$, SO_2, and VOC emissions, and interaction terms. Each of these emissions terms was included in an interaction with the PS* as well. Temperature, relative humidity, and the daily temperature and relative humidity-dependent dissociation constant for nitrate (Mozurkewich 1993) were also included. Model selection proceeded for the PM2.5 model in a way similar to that for the O3 model, i.e., covariates were removed one-by-one in order of decreasing significance. The final formulation is:

$$[PM_{2.5}] = ß_0 + (ß_1 PS^*)E_{EGU}^{NO_x} + (ß_2 PS^*)E_{EGU,reg}^{NO_x} + (ß_3 PS^*)E_{EGU}^{SO_2} + (ß_4 + ß_5 PS^*)E_{EGU,reg}^{SO_2} + (ß_6)E_{MOB}^{PM_{2.5}} + ß_8 WS + ß_9 Temp + ß_{10} RH + ß_{11} RF + \varepsilon_{PM_{2.5}} \quad \text{(Eq. 5.2)}$$

The statistical models (Eq. 5.1 and Eq. 5.2) were used to address the two goals stated previously (to determine the sensitivity of air pollutant concentrations to emissions changes and to use these sensitivities to calculate daily counterfactual concentrations). The models relate observations to emissions on the same day. Multi-day impacts are captured to an extent using the meteorology variables and PS*, but are difficult to estimate directly because of the correlation between consecutive days in emissions. Sensitivities of

pollutants are represented by the ß's in each equation. These, when multiplied by their respective covariate, yield the contribution to the concentration by each model input.

Estimated emissions were replaced with counterfactual emissions—described below—to estimate counterfactual concentrations. All other model inputs, including PS^* and ε, remain unchanged, since these are independent of emissions. The supplemental information includes an assessment of the relationships between model residuals and inputs. Both the O_3 and $PM_{2.5}$ models capture the variability at the middle-and lower quantiles, but the models have some difficulty fully capturing very high observations. These very polluted days are due to factors that are difficult to control for in a statistical model, such as specific combinations of meteorological factors or impacts of wildfire plumes.

5.3.4 Chemical transport model quantification of ozone sensitivities with CMAQ

It is of interest to compare empirical sensitivities based on measured ambient concentrations to alternative source apportionment techniques. The CMAQ model (Byun and Schere 2006) provides a detailed characterization of physics and chemistry governing the transport, removal and formation of air pollutants in the ambient air. In this study, results from a previous study (Liao et al. 2008) and newer results using CMAQ runs with updated parameters and smaller grid size were compared to the results obtained from the statistical O_3 model. CMAQ-DDM version 4.3 with SAPRC-99 chemical mechanism was applied in 2001 on a 36-km grid, and version 5.0.2 with CB05 chemical mechanism was applied on a 12-km grid. Such models have previously been used to simulate responses (or sensitivities) of ambient O_3 concentrations to changes in emissions of their precursors (Cohan et al. 2005; Dunker 1981; Yang et al. 1997). The 36 km CMAQ run was driven using results from the Fifth-Generation NCAR/Penn State Mesoscale Meteorological Model (MM5) (Grell et al. 1994; Seaman 2000), and the 12 km CMAQ used the Weather Research Forecast (WRF) model version 3.6.1. Values for the 36 km run were modeled for a 2001 climatological year (the meteorology was developed by down-scaling from a climate model), and the 12 km results and empirical values that are plotted are also from this year. Both processed precursor emissions using the Sparse Matrix Operator Kernel Emissions (SMOKE) (Houyoux et al. 2000).

We chose 2001 for the modeling episode both because it came before many of the large changes in emissions and because we have modeling results for both 36 and 12 km resolutions. Detrending results showed that meteorological variability had a slight negative effect (3.2 ppb) on summertime ozone concentrations, and little impact in the winter (Henneman et al. 2015).

CMAQ-DDM directly calculates the semi-normalized first-order (or linear) sensitivities of both gas- and condensed-phase pollutants to precursor emissions (Cohan et al. 2005; Napelenok et al. 2006), i.e., the semi-normalized first-order sensitivity ($S_{i,j}$) of pollutant concentration i (C_i) to source emissions j (E_j) is determined, effectively, as Yang et al. (1997):

$$S_{i,j} = \frac{\delta C_i}{\delta \alpha_j} \qquad \text{(Eq. 5.3)}$$

where α_j is the relative level of the emissions from source j base calculation, and has a nominal value of 1 (Cohan et al. 2005; Hakami et al. 2004). The sensitivities, as presented here, have the same units as the corresponding pollutant concentrations. These sensitivities are local and represent how pollutant concentrations respond to precursor emission changes if the systems were linear. It is recognized that the system is not linear, but numerous studies suggest the first-order (linear) response is accurate up to domain-wide emission changes of the order of 30–50% (depending on species) (Cohan et al. 2005; Hakami et al. 2003, 2004). The modeled concentrations and sensitivities for the 36 km run were first published in Liao et al. (2008).

5.3.5 Estimating emissions changes

Estimates of changes in emissions due to regulatory programs must take into account a number of factors that depend on the emission source. For EGUs, population growth, plant efficiency improvements, control installation dates, economic growth and decline, fuel type, etc. must be considered. On the other hand, vehicle fleet age and turnover, fuel type, and population are the important variables to consider for mobile source emissions.

Counterfactual emissions were calculated using a method similar to that used by Gégo et al. (2007). Their method takes into account all controls over the time period while correcting for changes in demand due to population growth or decline, economic trends,

etc. An annual emissions rate was calculated for the base year (BY) as the average ratio of daily emissions (tons (US)) to daily load (kilowatt hours— kWh).

$$ER_{BY} = \langle \frac{E(d)}{L(d)} \rangle_{BY} \quad \text{(Eq. 5.4)}$$

where L denotes load and d indexes day. ER_{BY} was assumed to remain constant for the counterfactual scenario of no controls. Annual counterfactual emissions were calculated by multiplying each day's load by ER_{BY}:

$$\mathrm{E}^{count}(\mathrm{d}) = \mathrm{ER}_{BY} * \mathrm{L}(\mathrm{d}, \mathrm{y}) \quad \text{(Eq. 5.5)}$$

where y is each year between BY and 2013. Complete CEM data was available for the ANAA starting in 1995, but data from multiple plants in the region were missing in 1995 and 1996, so 1997 is used as the BY for regional emissions.

This approach takes into account factors that cause changes in the emissions rate, e.g., controls, improvements in transmission efficiency, changing fuel costs that incentivize switching fuel type, etc. The model assumes that the application of controls did not differentially change plant dispatch. Another way to think about this is that the electricity demand in Atlanta is represented by the load carried by all of the plants in the area combined, which may not be the case if electricity is imported from or exported to plants in surrounding regions. This assumption is addressed by including regional emissions as a separate term in each regression equation, however, the limitation means that all deviations from observed concentrations calculated in the counterfactual concentrations below cannot be linked exclusively to specific controls. Comparison of counterfactual concentrations and dates when specific controls were installed allow for interpretation of emissions changes as attributable to specific regulatory actions and controls (details are discussed in Section "Counterfactual emissions").

Counterfactual mobile emissions were estimated using the "Rate of Progress" option in MOVES2010b, which models a scenario with no Clean Air Act Amendments by applying 1993 emission rates to all vehicles after this year (1993 is the default option for this scenario in MOVES) (USEPA 2012b). "Rate of Progress" still uses the same changes in vehicle fleet composition, vehicle miles traveled, and fuel formulations, but assumes 1993 emissions factors for new vehicles.

5.4 Results

5.4.1 Counterfactual emissions estimates

Differences between actual and counterfactual EGU NO_X emissions (Figure 5-3) show the largest reductions occur in 2002 and 2009, aligning with the beginning of summertime- specific NO_X controls (early 2000s) and a shift to year-round controls (late 2000s). The largest SO_2 emissions reductions occurred in 2009—when two large coal plants in the region had completed the installation of their flu gas desulphurization (FGD) technologies—and 2012—when much of the electricity load in Atlanta switched from coal to natural gas.

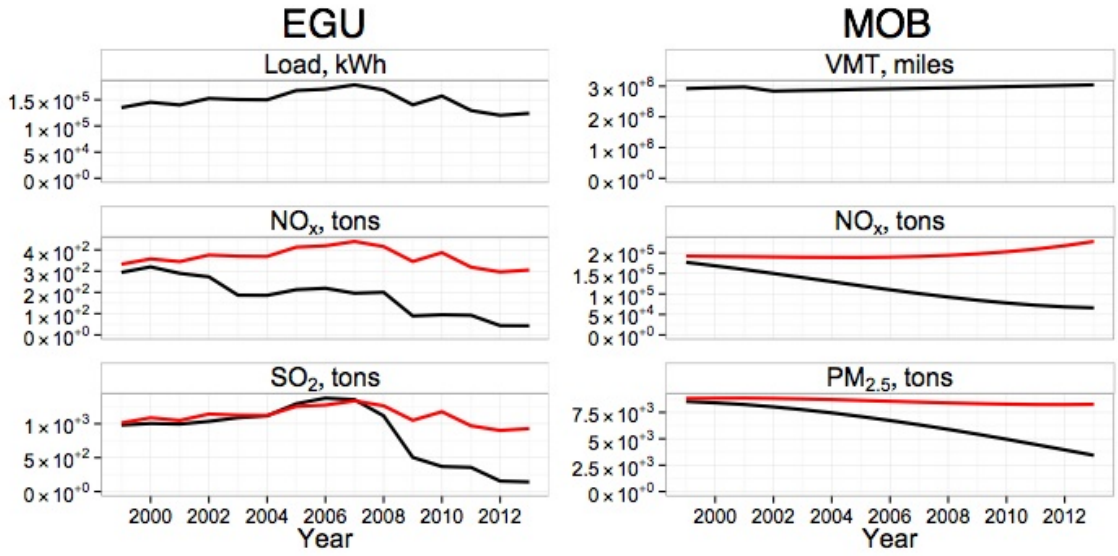


Figure 5-3. Left, actual (black) and counterfactual (red) emissions from EGUs. Right, actual and counterfactual emissions from mobile sources (MOB). The top graph in each column (load for EGUs and VMT (Vehicle Miles Traveled) for MOB) are the measure used for demand for each source category.

Both mobile NO_X and $PM_{2.5}$ emissions (Figure 5-3) show decreases between 2000 and 2012, corroborating the findings of Vijayaraghavan et al. (2014). Mobile source SO_2 emissions decrease dramatically (by 80% between 2004 and 2006) after fuels with reduced sulfur content were required year-round beginning in 2004. Other emitted species (e.g., VOCs, primary PM species) used in this study besides those plotted in Figure 5-3 generally follow the trend of NOX emissions. Counterfactual estimates show slightly increasing NO_X and $PM_{2.5}$ emissions that do not vary much between years and follow the estimated VMT.

5.4.2 *Empirical model evaluations*

The regression analysis led to a statistical model for O_3 (Table 5-2) with an R^2 and root mean square error (RMSE) of 0.67 and 11 ppb, respectively. The related values for the $PM_{2.5}$ model (Table 5-3) are 0.42 and 5.7 μg m^{-3}. Mean observed O_3 and $PM_{2.5}$ over this time period were 41 ppb and 14 μg m^{-3}, respectively. While the RMSE values are somewhat large, it is more revealing to compare the regression parameter magnitudes to their standard errors and their average, minimum, and maximum effect sizes. Standard errors are generally small compared to their coefficients in both the O_3 and $PM_{2.5}$ models, which suggests that the model covariates are predictive of changes in the response. Further, although many covariates have an average effect of zero (due to the scaling of PS^* and the meteorological covariates), many of their individual daily contributions are often of the same order of magnitude or larger compared to the intercept.

Table 5-2. Regression coefficients summary for Eq. 5.1 for O_3. Values on the left-hand side of the table represent values from the regression, and values on the right-hand ride summarize the total empirical sensitivities of ozone to emissions. The regression was performed on 4030 observations.

Covariate	Coefficient	Unit	Estimate	Std. Error	Average contribution [*ppb*]	Minimum contribution [*ppb*]	Maximum contribution [*ppb*]
Intercept	β_0	ppb	42	$5\text{x}10^{-1}$	42	–	–
$E^{NO_x}_{EGU}$	β_1	$ppb\,ton^{-1}$	$-4.1\text{x}10^{-3}$	$3\text{x}10^{-3}$	−0.74	−2	0.02
$E^{NO_x}_{EGU,reg}$	β_2	$ppb\,ton^{-1}$	$1.7\text{x}10^{-3}$	$8\text{x}10^{-4}$	2.1	0.46	5.3
$E^{NO_x}_{EGU,reg} * PS^*$	β_3	ton^{-1}	$1.1\text{x}10^{-4}$	$4\text{x}10^{-5}$	0.0	−9.7	19
$E^{NO_x}_{MOB}$	β_4	$ppb\,ton^{-1}$	$-4.2\text{x}10^{-3}$	$3\text{x}10^{-3}$	−1.3	−2.3	−0.41
$E^{NO_x}_{MOB} * PS^*$	β_5	ton^{-1}	$3.4\text{x}10^{-3}$	$2\text{x}10^{-4}$	0.0	−58	110
$E^{NO_x}_{MOB} * E^{MOB}_{VOC} * PS^*$	β_6	ton^{-2}	$-8.3\text{x}10^{-6}$	$1\text{x}10^{-6}$	0.0	−66	33
WS	β_7	$^1 ppb\,\|\|m\,s^{-1}\|\|^{-1}$	$-3.3\text{x}10^{-1}$	$2\text{x}10^{-1}$	0.0	−1.8	0.44
$Temp$	β_8	$^1 ppb\,\|\|°C\|\|^{-1}$	$3.5\text{x}10^{0}$	$3\text{x}10^{-1}$	0.0	−11	7.5
RH	β_9	$^1 ppb\,\|\|\%\|\|^{-1}$	$-1.7\text{x}10^{0}$	$2\text{x}10^{-1}$	0.0	−3.9	4.0
RF	β_{10}	$^{1,2} ppb\,\|\|Y\|\|^{-1}$	$8.6\text{x}10^{-2}$	$2\text{x}10^{-1}$	0.0	−0.06	0.13

[1] $\|\|unit\|\|$ denotes scaled and normalized by subtracting the mean and dividing by standard deviation

[2] Y denotes Rainfall is a factor (1-0) variable

Table 5-3. Regression coefficients summary for Eq. 5.2 for $PM_{2.5}$. Values on the left-hand side of the table represent values from the regression, and values on the right hand side summarize the total empirical sensitivities of $PM_{2.5}$ to emissions. The regression was performed on 3616 observations.

Covariate	Coefficient	Unit	Estimate	Std. Error	Average Contribution [$\mu g\, m^{-3}$]	Minimum Contribution [$\mu g\, m^{-3}$]	Maximum Contribution [$\mu g\, m^{-3}$]
Intercept	β_0	$\mu g\, m^{-3}$	5.5	$3x10^{-1}$	5.5	–	–
$E_{EGU}^{NO_x} * PS^*$	β_1	$\mu g\, m^{-3}\, ppb^{-1}\, ton^{-1}$	$-1.9x10^{-4}$	$1x10^{-4}$	0.0	−4.7	2.7
$E_{EGU,reg}^{NO_x} * PS^*$	β_2	$\mu g\, m^{-3}\, ppb^{-1}\, ton^{-1}$	$-5.3x10^{-5}$	$4x10^{-5}$	0.0	−9.4	4.9
$E_{EGU}^{SO_2} * PS^*$	β_3	$\mu g\, m^{-3}\, ppb^{-1}\, ton^{-1}$	$1.1x10^{-4}$	$3x10^{-5}$	0.0	−6.8	9.3
$E_{EGU,reg}^{SO_2}$	β_4	$\mu g\, m^{-3}\, ton^{-1}$	$1.5x10^{-3}$	$1x10^{-4}$	5.2	0.69	11
$E_{EGU,reg}^{SO_2} * PS^*$	β_5	$\mu g\, m^{-3}\, ppb^{-1}\, ton^{-1}$	$4.3x10^{-5}$	$2x10^{-5}$	0.0	−10	17
$E_{MOB}^{PM_{2.5}}$	β_5	$\mu g\, m^{-3}\, ton^{-1}$	$2.1x10^{-1}$	$3x10^{-2}$	3.7	0.81	5.8
$E_{MOB}^{VOC} * PS^*$	β_6	$\mu g\, m^{-3}\, ppb^{-1}\, ton^{-1}$	$4.3x10^{-4}$	$1x10^{-4}$	0.0	−6.4	3.7
WS	β_7	$^1 ppb\, \|\|m\, s^{-1}\|\|^{-1}$	$-1.7x10^{0}$	$1x10^{-1}$	0.0	−9.1	3.6
$Temp$	β_8	$^1 ppb\, \|\|°C\|\|^{-1}$	$1.3x10^{0}$	$2x10^{-1}$	0.0	−3.8	2.7
RH	β_9	$^1 ppb\, \|\|\%\|\|^{-1}$	$-2.0x10^{1}$	$1x10^{-1}$	0.0	−0.45	0.28
RF	β_{10}	$^{1,2} ppb\, \|\|Y\|\|^{-1}$	$-6.5x10^{1}$	$1x10^{-1}$	0.0	−1.0	1.4

[1] $\|\|unit\|\|$ denotes scaled and normalized by subtracting the mean and dividing by standard deviation

[2] Y denotes Rainfall is a factor (1-0) variable

Average contributions (Table 5-2 and Table 5-3), calculated by multiplying each regression coefficient by the average of the corresponding covariate, are a measure of the relative importance of each term in the regression. Plots of the daily contribution summed by source category (Figure 5-4) show that each source contribution varies by seasons. The intercepts for O_3 (42 ppb) and $PM_{2.5}$ (5.0 µg m^{-3}) are estimates of the average background concentrations in Atlanta over the study period that would occur without local and regional emissions from mobile and EGU sources, but would include long-range transport and the impact of other sources, and they include the average of all nonlinear responses to emissions not captured in the remaining terms.

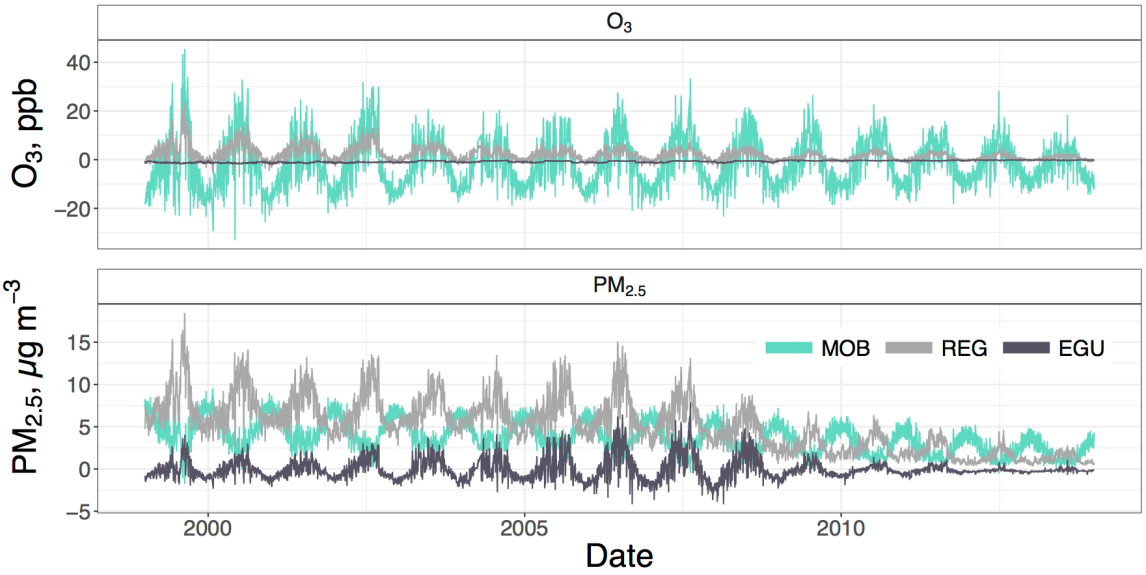


Figure 5-4. Sensitivities of daily O_3 and $PM_{2.5}$ concentrations to mobile (MOB), local EGU, and regional EGU (REG) emissions.

5.4.3 CMAQ-DDM model evaluation

CMAQ-modeled ozone concentrations from each model are from single grid cells that cover downtown Atlanta (including JST). Concentrations from both models exhibit similar annual and daily variability as observed concentrations (Figure S-6). Evaluation statistics for all days and days with observed O_3 over 60 ppb (Table 5-4) are somewhat higher than typical statistics reported in the meta-analysis of CTM results published by Simon et al. (2012). However, the current evaluation (i.e., at a single monitor) is stringent in comparison to others, which typically use many monitors at multiple locations. Between the two model runs, the 36 km model was biased higher than the 12 km, and correlation was higher for the 12 km. The overall high bias of the 36 km model improves the results for days over 60 ppb, though the correlation on these days is lower than for the 12 km model.

Table 5-4. Evaluation statistics for 2 versions of CMAQ, including the number of comparisons (N), normalized mean bias (NMB), normalized mean error (NME), mean bias (MB), and correlation (r).

CMAQ Model	N	NMB	NME	MB	r
36-km	356	38.24	48.78	16.18	0.54
12-km	361	-0.82	27.28	-0.35	0.70
36-km (>60)	80	6.60	23.76	4.79	0.14
12-km (>60)	80	-16.83	19.71	-12.23	0.43

Sensitivities of O_3 to EGU and mobile emissions in 2001 peak in the summertime and are negative in the winter and fall (Figure S-6). In the winter and spring, sensitivities produced by the CMAQ models agree more with each other than the empirical, and, in the summer the 12 -km model results agree more with the empirical results. 36 km sensitivities are biased high compared to the other models in the summertime.

While a direct comparison between observed and measured concentrations and sensitivities is important for putting results in perspective with other studies, this study focuses attention in the discussion on the comparison between sensitivities relative to O_3 levels. The goal of this analysis is to assess model intermediates and, relating to model outputs, somewhat reduces the impact of bias and differences in model inputs between model setups.

5.4.4 Counterfactual concentration estimates

Year-specific box plots of the actual and counterfactual ozone time series (Figure 5-5) show that median ozone values are relatively insensitive to emissions changes in EGUs and mobile sources (the observed median is 39.5 ppb and counterfactual median is 38.8 ppb). The bulk of the difference between actual and counterfactual appears in the highest and lowest ozone concentrations. As emissions have decreased, variability in annual ozone distributions has shrunk. Counterfactual emissions would have led to both more low-ozone days (days with $MA8hO_3$ below 30 ppb) and high ozone days. For example, 99 days were observed with $MA8hO_3$ below 30 ppb at JST in 2013, and the counterfactual estimates it would have been 134 days without controls. Four days were observed above 70 ppb (the recently promulgated O_3 NAAQS), whereas the counterfactual predicts 27. For 75 ppb, the similar values are one and 16 days, respectively. Differences

between observed and counterfactuals are small at the beginning of the time series, and increase as the differences between actual and counterfactual emissions grow.

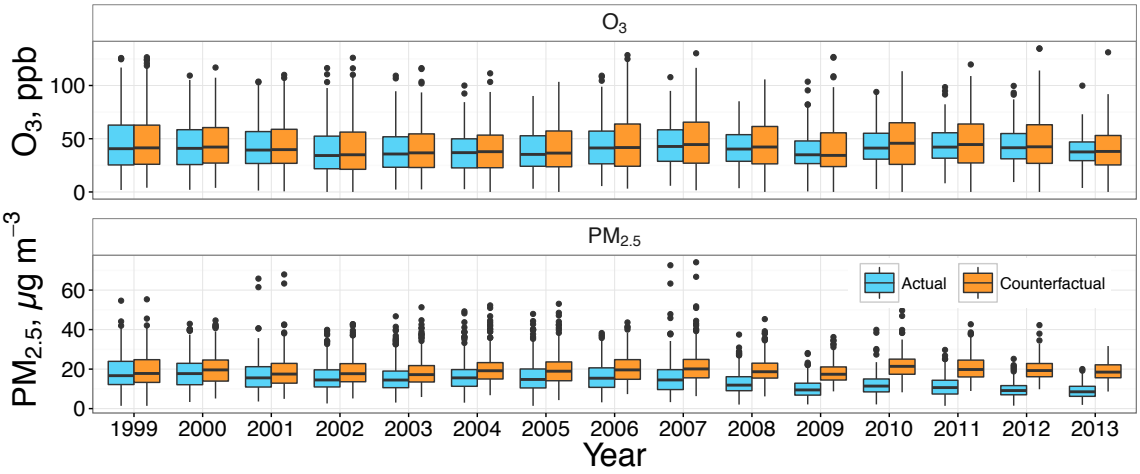


Figure 5-5. Annual box plot of the actual and counterfactual ozone and $PM_{2.5}$. Center lines are the median, the boxes are the first and third quartiles, and whiskers extend to values within 1.5 times the interquartile range.

5.5 Discussion

5.5.1 Counterfactual emissions

Differences between actual and counterfactual emissions (Figure 5-3) align well with known regulations and the resulting controls. Mobile emissions decrease in a near-linear fashion as old vehicles are replaced with new, cleaner ones, while VMT in the region has grown slowly. EGU emissions change more abruptly as controls are installed, plants change fuel types, and load is shifted between facilities. Information on when controls were installed and new plants were brought online were used to analyze changes in emissions. The information in this section has been taken from information available through EPA's Air Markets Database (USEPA 2013). The discussion focuses on sources within the ANAA (Figure 5-1).

Of the 12 EGUs in the ANAA (Figure 5-1), some are larger base load plants and others are smaller peaking plants. One plant near downtown Atlanta, McDonough, was converted from coal to natural gas between 2011 and 2012, and another, Wansley, began operation of 10 natural gas units between 2002 and 2012 while still keeping its coal units available. Similarly, all six of the smaller plants that have gone online since 1999 (Chattahoochee Energy Facility, Doyle Generating Facility, Hawk Road Energy Facility,

MPC Generating Facility, Tenaska Georgia Generating Station, and Walton County Power Facility) use natural gas. Plant Atkinson, a small plant near downtown, was run on natural gas between 1995 and 2006, when it was retired. The remaining plants (Bowen, Harllee Branch, and Yates) ran primarily on coal between 1995 and 2012. In the ANAA plants, 3% of the total load was generated by units with natural gas as the primary fuel in 2008 compared with 21% in 2012 (USEPA 2013), a change that is attributable to both regulations and reduced natural gas prices, and was planned many years in advance (Georgia Power 2007).

The load on the plants increased by 4% between 1995 and 2012, with a peak in 2007 (Figure 5-3). Most plants have added low NO_X burners, and some (Bowen, Chattahoochee Energy Facility, and Wansley) have installed selective catalytic reduction (SCR) NOx controls. Annual NO_X emissions by the 12 local plants fell by 85% between 1995 and 2012. At first, regulations led to the SCRs being operated only during the summer months (May-September). Beginning in 2009, NO_X controls on the largest plants in the region were operated year-round.

Years with the greatest increase in the difference between actual and counterfactual EGU NO_X emissions (Figure 5-1) are 2002–2003 and 2008–2009. Between 2002 and 2003, plants Bowen and Wansley completed installation of SCR NO_X controls, the Chattahoochee Energy Facility, which was built with SCR technology, went online, and Harllee Branch installed low NO_X burners.

Plants Bowen and Wansley installed FGD technologies on their coal units in 2008 and 2009. These controls, along with the switch to natural gas, have contributed to a decrease in SO_2 emissions by EGUs in Atlanta of 81% between 1995 and 2012. The years between 2008 and 2010 saw the greatest decrease in SO_2 emissions compared to the counterfactual. Nearly all of this decrease can be traced to the installation of FGD controls at plant Bowen and Wansley. Further reductions were achieved by relying less on coal-fired plants, including the conversion of Plant McDonough from coal to natural gas in 2012.

In Georgia, because the cost of controls cannot be recovered by raising electricity rates unless the control is deemed necessary under existing law, it is assumed that all controls can be attributed to regulatory actions, and these could potentially be future

anticipated regulations (Georgia Power 2007). Fuel switches, plant commissioning and retirement schedules, and electricity trading between utilities are governed by complex relationships that include current/projected fuel prices, varying costs of producing electricity at different plants, projected demand, and anticipated future regulatory actions. Therefore, while the estimated emissions reductions are tied to controls, not all emissions changes calculated can be attributed exclusively to regulatory actions.

5.5.2 Magnitudes of EGU and mobile sensitivities

Parameters from the models in Eq. 5.1 and Eq. 5.2 (Table 5-2 and Table 5-3) provide the relative importance of each source-pollutant contribution to concentrations. It is important to recognize that modeled mobile emissions are highly correlated across species. Therefore, it is difficult to separate the effects of all species of interest in a statistical model. For the O_3 model, VOC emissions were included, but only in interactions terms because of this cross-species co-linearity. Mobile source carbon monoxide (CO) emissions, which contribute to O_3 formulation along with VOCs and NO_X, are highly correlated with VOC emissions, so the sensitivity to VOC emissions includes the impacts from CO. Contributions of mobile emissions to total O_3 and $PM_{2.5}$ concentrations are best interpreted as a sum of the component emissions (Figure 5-4).

Modeled mobile contributions to ozone concentrations are dominated by the interactions between NO_X emissions and PS^* and NO_X and VOC emissions and PS^*. Average contributions of both of these is zero because the terms are normalized to zero, but the magnitudes of the maxima and minima are large compared to the other terms, i.e., NO_X emissions lead to high ozone on photochemically-active days, but reductions when the meteorology is not conducive to O_3 formation. The negative coefficient on the interaction term that includes VOC emissions and PS^* shows the importance of VOC-limited conditions when increased VOC emissions lead to increases in otherwise low ozone concentrations, i.e., during radical-limited periods there is a positive sensitivity to VOC emissions and negative sensitivity to NO_X emissions. This happens during most of the non-summer season as well as on lower O_3 days during the summer.

For 13 monitors within 74 km of JST, correlations (Pearson R) were at least 0.77 for $MDA8hO_3$ from 2002-2010, and measurements show consistent annual trends at urban, suburban and rural sites (Figure S-1). NO_X concentrations, however, show greater spatial

variability; therefore, concentration sensitivities to emissions estimated at JST may differ across the region.

The sensitivities of ozone to emissions (Figure 5-4) suggest that mobile emissions have a greater effect on ozone levels at JST than EGU emissions. As annual ozone distributions have shrunk since 2000, the sensitivities have also decreased in magnitude. Tong et al. (2006), Muller et al. (2009) and others have found ground-level NO_X emissions have a much greater impact on ozone concentrations than stack emissions from outside of the city.

All EGU SO_2 emissions terms are of particular interest in the $PM_{2.5}$ model due to the importance of sulfate in Atlanta. The sum of the mean contributions of these terms is 5.2 μg m^{-3}, which corresponds to 36% of the average total $PM_{2.5}$ across all days (14.2 μg m^{-3}). Over the same time period, sulfate, a secondary particulate species mainly attributable to atmospheric processing of EGU SO_2 emissions, made up 27% of measured $PM_{2.5}$ at JST. Ammonium particulate matter is strongly associated with sulfate (sulfuric acid will react with ammonia gas), and recent results find that sulfate can enhance biogenic secondary organic aerosol formation (Marais et al. 2016; Weber et al. 2016), explaining when the sensitivity to SO_2 emissions is greater than the measured sulfate.

Mobile sources are estimated to have contributed an average of 3.7 μg m^{-3} (26%). The measured species important to the total over this time period are organic carbon (28%), elemental carbon (11%), ammonium (11%), and nitrate (6%). Each of these remaining species is associated with mobile emissions, and may be a portion of the 26% they contribute. As in the ozone model, it is most appropriate to interpret the contributions from mobile emissions as the sum of their parts instead of by individual species. While major contributors (primary $PM_{2.5}$ and NO_X) are included in the model, their co-linearity with other species emissions means that the total captured is likely the joint effect of all mobile emissions emissions. For example, primary $PM_{2.5}$ emissions are highly co-linear with VOC emissions (VOC emissions are still included in the model in an interaction with PS^*).

Two terms of interest in the $PM_{2.5}$ model are the interactions between EGU and REG NOx emissions and PS^*. These terms are negative, meaning that increasing emissions correspond with increasing $PM_{2.5}$ levels in the winter- time, and a negative contribution in the summer. Positive contributions in the winter correspond with increased nitrate levels,

a secondary species that forms when NO_X is oxidized to HNO_3, which then reacts with NH_3 to form ammonium nitrate. Brock et al. (2002) showed that young NO_X plumes decrease the conversion of SO_2 to sulfate because of decreased radicals due to NOx titration, but cautioned that this effect is generally assumed small and uncertain. The current study finds evidence of this phenomenon, and attributes a mean daily reduction of 1 μg m^{-3} in summers (May-September) across the time period, though ins importance has decreased over time as both NO_X and SO_2 emissions have been reduced. Total $PM_{2.5}$ sensitivities to EGU emissions are highest in the summer, which corresponds with both increased SO_2 emissions and increased photochemical activity that contributes to elevated sulfate concentrations. Mobile sensitivities exhibit much less annual variability than EGUs, and contributions from both source categories have decreased over time as emissions have decreased.

5.5.3 Counterfactual concentrations

The largest changes between observed and counterfactual concentrations occur after 2009, when control programs had at or near their greatest impact on emissions. Between 2010 and 2013, Atlanta experienced many fewer high-O_3 and $PM_{2.5}$ days than it would have without regulations (Table 5-5). Values for comparison in Table 5-5 were chosen based on standards of regulatory importance. For example, 12.0 and 15.0 μg m^{-3} are the primary and secondary annual mean NAAQS for $PM_{2.5}$, and 35 μg m^{-3} is the primary and secondary 24-hr NAAQS. The O_3 NAAQS was recently changed from 75 ppb to 70 ppb, though the lower end of the proposed range was 60 ppb (U.S. EPA 2014). Results here show that regulatory programs have had important influence on the concentrations of regulatory importance in Atlanta.

Table 5-5. The number of days on which observed and counterfactual O_3 and $PM_{2.5}$ were greater than concentrations of regulatory importance from 2010 to 2013, when regulatory policies were at or near their greatest impact.

	Observed	**Counterfactual**
O_3 > 60 ppb	200	396
O_3 > 60 ppb	75	223
O_3 > 60 ppb	42	164
$PM_{2.5}$ > 12.0 μg m^{-3}	456	1391
$PM_{2.5}$ > 15.0 μg m^{-3}	222	1164
$PM_{2.5}$ > 35 μg m^{-3}	2	7

The Georgia Department of Natural Resources reports 46 exceedances of the 8-hr ozone standard in 2000 (http://www.air.dnr.state.ga.us/). The highest reported MDA8hr in that year in the ANAA was 139 ppb, which is 20 ppb higher than the highest value measured at JST that year. The maximum MDA8hr value observed at JST is 131 ppb (in the summer of 1998—the year that the station began recording ozone data). There are 4 days in which the counterfactual ozone exceeds 130 ppb, all of which occur in 2007 and later (Figure 5-5). The shapes of the annual distributions of counterfactual ozone in the later years resemble those of observed ozone early in the time period. Decreased emissions have had the effect of decreasing median concentrations and distribution widths for $PM_{2.5}$. In 2013, the observed median $PM_{2.5}$ is 8.9 μg m^{-3}, compared with a counterfactual of 18.9 μg m^{-3}. The largest reductions occurred in 2009 and onwards, coinciding with the years of greatest reductions in SO_2 emissions from EGUs.

For both O_3 and $PM_{2.5}$, EGU and REG NO_X sensitivities are positive in the summertime and negative in the wintertime. Therefore, summertime emissions reductions—for instance, NO_X emissions reductions occurring in the mid- 2000s—typically reduced concentrations of both pollutants. Winter sensitivities, however, are negative, meaning wintertime NO_X controls, in wider use in the late 2000s, likely increased concentrations. For $PM_{2.5}$, this effect is small compared to the effect of reduced SO_2 emissions.

5.5.4 Comparison to CMAQ-calculated sensitivities

As reported here, CMAQ-modeled ozone sensitivities rep- resent the first order change in ozone expected from a 100% reduction in anthropogenic emissions. Sensitivities show the contribution of the sources to the measured con- centration calculated in the modeled grid cell of interest, which, for this work, is the grid cell that corresponds to JST. CMAQ-modeled sensitivities provide a point of comparison for empirical sensitivities calculated using a different, independent method.

CMAQ-modeled and empirically-calculated sensitivities to all anthropogenic NO_X emissions have a positive relation- ship with O_3 concentrations (Figure 5-6). Empirical sensitivities show more variability and a greater magnitude across the same range than CMAQ-modeled sensitivities from both grid resolutions. From a regulatory perspective, the ozone concentration that corresponds to a sensitivity of zero is of interest—at ozone

concentrations above this point, NO_X controls reduce ozone; at concentrations below this point, controls increase ozone. The three models estimate a range of 16.8 ppb for this value (42.9 (95% CI 41.3–44.5), 57.9 (56.5–59.3), and 59.7 (57.2–62.6) ppb for empirical, 36 km CMAQ, and 12 km CMAQ, respectively). The slope of the empirical sensitivities (0.63 ppb ppb^{-1}) is slightly greater than that for CMAQ (0.48 ppb ppb^{-1}), leading to a closer agreement of the two models at higher O_3 concentrations. The use of a climatologic year will also influence the cross-over points.

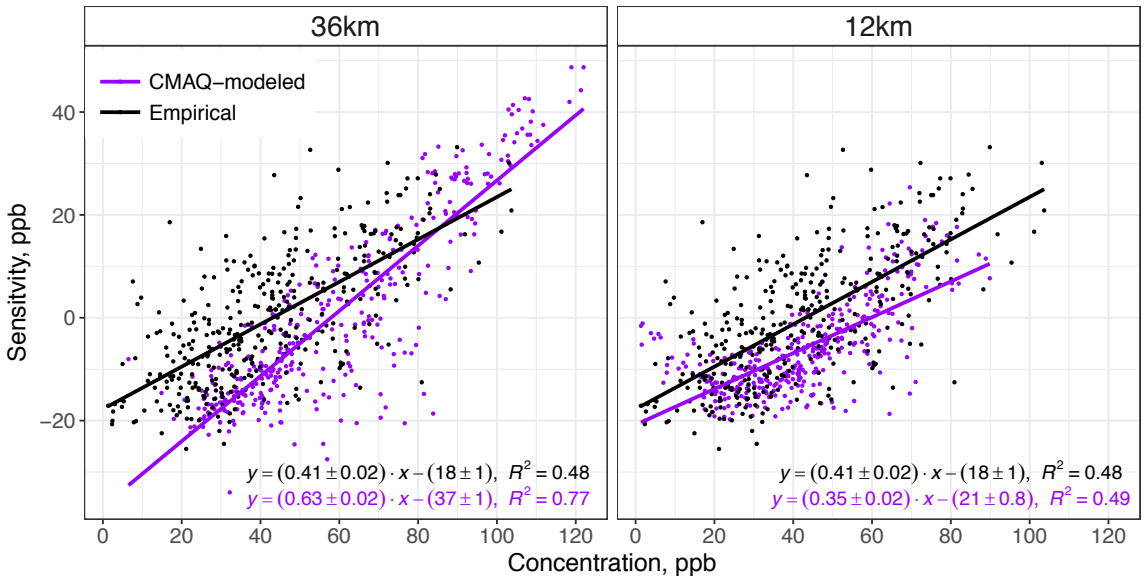


Figure 5-6. 2001 ozone sensitivities to anthropogenic NO_X emissions calculated by the empirical method and CMAQ for both the 36 km and 12 km runs. Sensitivities greater than zero occur in conditions where NO_X emissions increase ozone, and sensitivities less than zero indicate conditions where NO_X emissions reduce ozone. The x-axis intercepts of the lines are 42.9 (95% CI 41.3–44.5) for the empirical model, 57.9 (56.5–59.3) for 36 km CMAQ, and 59.7 (57.2–62.6) for 12 km CMAQ. Numbers in parentheses in the equations are the standard error of the regression coefficients.

Fewer modeled O_3 high (concentration greater than 80 ppb) days leads to a closer agreement between the empirical model and the 12 km CMAQ results than the 36 km CMAQ results. Recent evidence has shown that models may overestimate mobile NO_X emissions by as much as 50% (Anderson et al. 2014; Goldberg et al. 2016; Souri et al. 2016; Travis et al. 2016), which would generally decrease modeled ozone concentrations and sensitivities in the city center. This phenomenon may help explain the general underestimate of sensitivities in the CMAQ results.

Differences between empirical and CMAQ-modeled sensitivities are due, in part, to differences in emissions inputs used in the models. Although the empirical sensitivities are estimated using emissions only in the Atlanta area, long-term regional emissions trends are similar due to national regulatory actions impacting nearby states during similar timeframes. Besides local contributions, empirical sensitivities are impacted by regional emissions, and correlations between the two may impact the results. CMAQ sensitivities are estimated for emissions within the entire modeling domain.

Cohan et al. (2005) investigated the magnitudes of first- and second-order CMAQ-DDM sensitivities of ozone in Atlanta to NO_X and VOC emissions. They found that second-order sensitivities are of similar magnitude to first order sensitivities on high-ozone days, which may further explain the difference between empirical and CMAQ- estimated sensitivities in the present study—CMAQ is used to calculate only first-order sensitivities, while the statistical model captures higher order sensitivities because of the PS^* and nonlinear term.

5.5.5 Method limitations

While this work shows benefits of statistical and numerical air quality modeling, both approaches have limitations. Statistical modeling is subject to bias as a result of correlated input variables, confounding, model selection, and data errors, including estimated emissions. The current study, in particular, is limited by modeled mobile emissions; each of the species is highly correlated with the others, and daily variability is dominated by the approximation of typical weekday/weekend differences. The high level of correlation between species limits the interpretation of ambient sensitivities to mobile emissions—the combined effects from a specific source are more relatable to known physical processes than individual source-species terms in the model.

The use of PS^*, an approximation for emissions-independent photochemical oxidative state, allows for the estimation of emissions-meteorology relationships. However, some meteorological conditions, such as days with wind from specific sources, are not fully captured by this method. Attempts to split daily emissions inputs by wind direction did not improve the predictive ability, and produced results inconsistent with established atmospheric relationships.

Counterfactual emissions assume that changes in electricity production by ANAA EGUs are independent of regulations, and that deviations from the historical emissions-load relationship can be attributed to regulatory actions. Both assumptions are limitations of the method, but comparisons of observed changes in this relationship and regulatory implementation dates provide evidence that the assumptions are reasonable.

Daily counterfactual concentrations, which maintain the autocorrelation of observed ambient concentrations, are limited by the variability in measurements, particularly in the $PM_{2.5}$ model. At the beginning of the time series, observations showed high scatter around the annual mean, which the statistical model is proficient at estimating. The scatter is smaller towards the end of the time series, therefore, counterfactual estimates in the later years lack the very high days observed in the early part of the time series (Figure 5-5). Further, in Atlanta, high $PM_{2.5}$ days can be associated with wild land fires, which are not studied here.

5.6 Conclusion

We presented a detailed accountability assessment of regulatory impacts on O_3 and $PM_{2.5}$ concentrations in Atlanta, GA between 1999 and 2013. The atypical approach addressed challenges typical in accountability programs. Applying emissions factors to create counterfactual concentration and relating these to a detailed regulatory assessment increased confidence in the changes attributed to regulations, even as regulations were implemented incrementally over time.

The empirical method that employs statistical modeling to develop daily sensitivities and counterfactuals of measured pollutants to changes in emissions. Empirical relationships in the model are reliant on PS^*, a daily metric for the photochemical state of the atmosphere that varies with temperature, relative humidity, wind speed, and rainfall. The model was applied to a central monitoring site in Atlanta, GA using estimated EGU and mobile emissions from the entire metro Atlanta area. Empirical sensitivities agree with sensitivities estimated using CMAQ-DDM, a CTM that explicitly accounts for atmospheric chemistry and physics. The comparison between model intermediates across platforms provides evidence that the empirical relationships appropriately capture emissions relationships with observed ambient concentrations.

For this monitoring location, sensitivities of ozone to mobile NO_X and VOC emissions dominate those of EGU NO_X emissions in magnitude, a result that may be different in locations outside of the city center. Minimum and maximum annual contributions of both mobile and EGU emissions have decreased over time with the implementation of controls.

Mobile emissions dominate the contribution to PM2.5 concentrations in the winter, and EGU SO2 emissions dominate in the summer. EGU contributions to measured PM2.5 concentrations show a large seasonal pattern, which has decreased significantly with the implementation of SO2 emissions controls. Average EGU contributions are similar to average sulfate concentrations in Atlanta, and regional and local EGU's contribute similar amounts to sulfate on average.

Statistical models of this type are relatively straightforward to apply, and daily counterfactuals estimated using these models are appropriate for use in accountability studies investigating public health responses to air quality regulations. This approach could give policy-makers a ready estimate of impacts of past or proposed controls. Then, in areas with the greatest interest (or to investigate specific effects or certain controls), policy-makers could employ more sophisticated CTMs to model changes.

5.7 Acknowledgements

This material is based upon work supported by Health Effects Institute and the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1148903. Detailed data and guidance was provided for MOVES modeling by Gil Grodzinsky and Jon Morton of the Georgia Environmental Protection Division Air Protection Branch. We thank Charles Huling, formerly of Southern Company, for his input.

5.8 References

Anderson DC, Loughner CP, Diskin G, Weinheimer A, Canty TP, Salawitch RJ, Worden HM, Fried A, Mikoviny T, Wisthaler A, Dickerson RR (2014) Measured and modeled CO and NOy in DISCOVER-AQ: An evaluation of emissions and chemistry over the eastern US. Atmos Environ 96:78–87. doi:10.1016/j.atmosenv.2014.07.004, http://linkinghub.elsevier.com/retrieve/pii/S1352231014005251

Appel KW, Gilliland AB, Sarwar G, Gilliam RC (2007) Evaluation of the Community Multiscale Air Quality (CMAQ) model version 4.5: Sensitivities impacting model performance. Atmos Environ 41(40):9603–9615. doi:10.1016/j.atmosenv.2007.08.044, http://linkinghub.elsevier.com/retrieve/pii/S1352231007007534

Blanchard C, Hidy G, Tanenbaum S (2010) NMOC, ozone, and organic aerosol in the southeastern United States, 19992007: 2. Ozone trends and sensitivity to NMOC emis- sions in Atlanta, Georgia. Atmos Environ 44(38):4840–4849. doi:10.1016/j.atmosenv.2010.07.030, http://linkinghub.elsevier.com/retrieve/pii/S1352231010005996

Blanchard CL, Hidy GM (2005) Effects of SO2 and NOx Emission Reductions on $PM_{2.5}$ Mass Concentrations in the South- eastern United States. J Air Waste Manag Assoc 55(3):265–272. doi:10.1080/10473289.2005.10464624

Bloomfield P, Royle JA, Steinberg LJ, Yang Q (1996) Account- ing for meteorological effects in measuring urban ozone levels and trends. Atmos Environ 30(17):3067–3077. doi:10.1016/1352-2310(95)00347-9, http://linkinghub.elsevier.com/retrieve/pii/1352231095003479

Brock CA, Washenfelder RA, Trainer M, Ryerson TB, Wilson JC, Reeves JM, Huey LG, Holloway JS, Parrish DD, Huebler G, Fehsenfeld FC (2002) Particle growth in the plumes of coal-fired power plants. J Geophys Res: Atmos 107(D12):AAC 9–1–AAC 9–14. doi:10.1029/2001JD001062

Byun DW, Schere K (2006) Review of the governing equations, computational algorithms, and other components of the Models- 3 Community Multscale Air Quality (CMAQ) modeling system. Appl Mech Rev pp 51–77

Byun DW, Kim ST, Kim SB (2007) Evaluation of air quality models for the simulation of a high ozone episode in the Houston metropolitan area. Atmos Environ 41(4):837–853. doi:10.1016/j.atmosenv.2006.08.038, http://linkinghub.elsevier.com/retrieve/pii/S1352231006008752

Camalier L, Cox W, Dolwick P (2007) The effects of meteorology on ozone in urban areas and their use in assessing ozone trends. Atmos Environ 41(33):7127–7137. doi:10.1016/j.atmosenv.2007.04.061, http://linkinghub.elsevier.com/retrieve/pii/S1352231007004165

Chang HH, Hao H, Sarnat SE (2014) A statistical modeling framework for projecting future ambient ozone and its health impact due to climate change. Atmos Environ 89:290–297. doi:10.1016/j.atmosenv.2014.02.037, http://linkinghub.elsevier.com/retrieve/pii/S1352231014001332

Cohan DS, Hakami A, Hu Y, Russell AG (2005) Nonlinear response of ozone to emissions: source apportionment and sensitivity analysis. Environ Sci Technol 39(17):6739–48. http://www.ncbi.nlm. nih.gov/pubmed/16190234

Davies T, Kelly P (1992) Surface Ozone Concentrations in Europe' Links With the Regional-Scale Atmospheric Circulation. J Geo- phys Res 97:9819–9832. http://onlinelibrary.wiley.com/doi/10.1029/92JD00419/full

Digar A, Cohan DS (2010) Efficient characterization of pollutant- emission response under parametric uncertainty. Environ Sci Technol 44(17):6724–30. doi:10.1021/es903743t, http://www.ncbi. nlm.nih.gov/pubmed/20701284

Dunker AM (1981) Efficient calculation of sensitivity coefficients for complex atmospheric models. Atmospheric Environment (1967) (1), http://www.sciencedirect.com/science/article/pii/ 000469818190305X

Dunker AM (1984) The decoupled direct method for calculating sensitivity coefficients in chemical kinetics. J Chem Phys 81(5):2385. doi:10.1063/1.447938

EPD (2009) Georgia's State Implementation Plan for the Atlanta 8- Hour Ozone Nonattainment Area

EPD (2012) Georgia's Redesignation Request and Maintenance Plan for the Atlanta Nonattainment Area for the 1997 $PM_{2.5}$ NAAQS

EPD (2014) Council regulation (EU) no 269/2014. http://eur- lex.europa.eu/legal-content/EN/TXT/?qid=1416170084502&uri=CELEX:32014R0269

Foley KM, Dolwick P, Hogrefe C, Simon H, Timin B, Possiel N (2015a) Dynamic evaluation of CMAQ part II: Evaluation of relative response factor metrics for ozone attainment demonstrations. Atmos Environ 103:188–195. doi:10.1016/j.atmosenv.2014.12.039, http://linkinghub.elsevier. com/retrieve/pii/S135223101400987X

Foley KM, Hogrefe C, Pouliot G, Possiel N, Roselle SJ, Simon H, Timin B (2015b) Dynamic evaluation of CMAQ part I: Separating the effects of changing emissions and changing meteorology on ozone levels between 2002 and 2005 in the eastern US. Atmos Environ 103(x):247–255. doi:10.1016/j.atmosenv.2014.12.038, http://linkinghub.elsevier.com/retrieve/pii/S1352231014009868

Garcia VC, Gego E, Lin S, Pantea C, Rappazzo K, Wootten A, Rao ST (2011) An evaluation of transported pollution and respiratory- related hospital admissions in the state of New York. Atmos Pollut Res 2(1):9–15. doi:10.5094/apr.2011.002

Ge´go E, Porter PS, Gilliland A, Rao ST (2007) Observation-Based Assessment of the Impact of Nitrogen Oxides Emissions Reductions on Ozone Air Quality over the Eastern United States. J Appl Meteorol Climatol 46(7):994–1008. doi:10.1175/JAM2523.1

Georgia EPD (2013) Rules for Air Quality Control. Tech. rep., Georgia EPD, http://www.georgiaair.org/airpermit/html/planningsupport/ naa.htm

Georgia Power (2007) 2007 Georgia Power Integrated Resources Plan Gilliland AB, Hogrefe C, Pinder RW, Godowitch JM, Foley KL, Rao S (2008) Dynamic evaluation of regional air quality models: Assessing changes in O3 stemming from changes in emissions and meteorology. Atmos Environ 42(20):5110–5123. doi:10.1016/j.atmosenv.2008.02.018, http://linkinghub.elsevier.com/retrieve/pii/S1352231008001374

Goldberg DL, Vinciguerra TP, Anderson DC, Hembeck L, Canty TP, Ehrman SH, Martins DK, Stauffer RM, Thompson AM, Salawitch RJ, Dickerson RR (2016) CAMx Ozone Source Attribution in the Eastern United States using Guidance from Observations during DISCOVER-AQ Maryland. Geophys Res Lett (43):2249–2258. doi:10.1002/2015GL067332

Grell G, Dudhia J, Stauffer D (1994) A description of the fifth generation Penn State/NCAR mesoscale model (MM5). NCAR Techni- cal Note, NCAR/TN-398+STR. National Center for Atmospheric Research, Boulder, Colorado

Hakami A, Odman M, Russell A (2003) Environmental science and technology. Atmos Environ 37:2442–2452

Hakami A, Odman T, Russell AG (2004) Nonlinearity in atmospheric response: A direct sensitivity analysis approach. J Geophys Res 109(D15):D15,303. doi:10.1029/2003JD004502

Hanna SR, Lu Z, Frey HC, Wheeler N, Vukovich J, Arunachalam S, Fernau M, Hansen DA (2001) Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain. Atmos Environ 35(5):891–903. doi:10.1016/S1352-2310(00)00367-8

Hansen DA, Edgerton ES, Hartsell BE, Jansen JJ, Kandasamy N, Hidy GM, Blanchard CL (2003) The southeastern aerosol research and characterization study: Part 1Overview. J Air Waste Manag Assoc 53(12):1460–1471. doi:10.1080/10473289.2003.1046 6318

Harrington W, Morgenstern R, Shih JS, Bell ML (2012) Did the Clean Air Act Amendments Of 1990 really improve air quality? Air Qual, Atmos Health 5(4):353–367. doi:10.1007/s11869-012-0176-5

HEI (2003) Assessing the health impact of air quality regulations: Concepts and methods for accountability research. Technical report, The Health Effects Institute

Henneman L, Holmes H, Russell A, Mullholand J (2015) Meteoro- logical Detrending of Primary and Secondary Pollutant Concentrations: Method Application and Evaluation Using Long-Term (2000-2012) Detailed Data. Atmospheric Environment

Houyoux M, Vukovich J, Coats C, Wheeler N, Kasibhatla P (2000) Emission inventory development and processing for the Seasonal Model for Regional Air Quality (SMRAQ) project. J Geophys Res: Atmos 105:9079–9090

Hu Y, Balachandran S, Pachon JE, Baek J, Ivey C, Holmes H, Odman MT, Mulholland JA, Russell AG (2014) Fine particulate matter source apportionment using a hybrid chemical transport and receptor model approach. Atmos Chem Phys 14(11):5415–5431. doi:10.5194/acp-14-5415-2014

Kuebler J, Van den Bergh H, Russell AG (2001) Long-term trends of primary and secondary pollutant concentrations in Switzerland and their response to emission controls and economic changes. Atmos Environ 35(8):1351–1363. doi:10.1016/S1352-2310(00)00401-5

Liao KJ, Tagaris E, Napelenok SL, Manomaiphiboon K, Woo JH, Amar P, He S, Russell AG (2008) Current and future linked responses of ozone and $PM_{2.5}$ to emission controls. Environ Sci Technol 42(13):4670–5. http://www.ncbi.nlm.nih.gov/pubmed/ 18677989

Marais EA, Jacob DJ, Jimenez JL, Campuzano-Jost P, Day DA, Hu W, Krechmer J, Zhu L, Kim PS, Miller CC, Fisher JA, Travis K, Yu K, Hanisco TF, Wolfe GM, Arkinson HL, Pye HOT, Froyd KD, Liao J, McNeill VF (2016) Aqueous-phase mechanism for secondary organic aerosol formation from isoprene: Application to the south- east United States and co-benefit of SO2 emission controls. Atmos Chem Phys 16(3):1603–1618. doi:10.5194/acp-16-1603-2016

Mozurkewich M (1993) The dissociation constant of ammonium nitrate and its dependence on temperature, relative humidity and particle size. Atmos Environ Part A Gen Top 27(2):261–270. doi:10.1016/0960-1686(93)90356-4

Muller N, Tong D, Mendelsohn RO (2009) Regulating NOx and SO2 Emissions in Atlanta. BE J Econ Anal Policy: Contrib Econ Anal Policy 9(2):. http://ideas.repec.org/a/bpj/bejeap/v9y2009i2n3.html

NAP (2004) Air Quality Management in the United States. The National Academies Press, Washington, DC, http://www.nap.edu/ catalog/10728/air- quality- management- in- the-united- states

Napelenok S, Cohan D, Hu Y, Russell A (2006) Decoupled direct 3D sensitivity analysis for particulate matter (DDM-3D/PM). Atmos Environ 40:6112–6121

Pope CA, Ezzati M, Dockery DW (2009) Fine-particulate air pollution and life expectancy in the United States. Engl J Med 360(4):376– 86. doi:10.1056/NEJMsa0805646, http://www.pubmedcentral. nih.gov/articlerender.fcgi?artid=3382057&tool=pmcentrez& rendertype=abstract

Porter PS, Rao ST, Hogrefe C, Gego E, Mathur R (2015) Methods for reducing biases and errors in regional photochemical model outputs for use in emission reduction and exposure assessments. Atmos Environ 112:178–188. doi:10.1016/j.atmosenv.2015.04.039, http://linkinghub.elsevier. com/retrieve/pii/S1352231015300455

R Core Team (2015) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/

Rich DQ, Kipen HM, Huang W, Wang G, Wang Y, Zhu P, Ohman- Strickland P, Hu M, Philipp C, Diehl SR, Lu SE, Tong J, Gong J, Thomas D, Zhu T, Zhang JJ (2012) Association between changes in air pollution levels during the Beijing Olympics and biomarkers of inflammation and thrombosis in healthy young adult. J Am Med Assoc 307(19):2068–2078. doi:10.1001/jama.2012.3488, NIHMS150003

Seaman N (2000) Fast, direct sensitivity analysis of multidimensional photochemical models. Atmos Environ 34:2231–2259

Seinfeld J, Pandis S (2006) Atmospheric Chemistry and Physics: From Air Pollution to Climate Change, 2nd edn. Wiley, New York

Simon H, Baker KR, Phillips S (2012) Compilation and interpretation of photochemical model performance statistics pub- lished between 2006 and 2012. Atmos Environ 61:124–139. doi:10.1016/j.atmosenv.2012.07.012

Souri AH, Choi Y, Jeon W, Li X, Pan S, Diao L, Westenbarger DA (2016) Constraining NOx emissions using satellite NO2 measurements during 2013 DISCOVER-AQ Texas campaign. Atmos Environ 131(2):371–381. doi:10.1016/j.atmosenv.2016.02.020

Tong DQ, Muller NZ, Mauzerall DL, Mendelsohn RO (2006) Policy Analysis Integrated Assessment of the Spatial Variability of Ozone Impacts from Emissions of Nitrogen Oxides. Environ Sci Technol 40(5):1395–1400

Travis KR, Jacob DJ, Fisher JA, Kim PS, Marais EA, Zhu L, Yu K, Miller CC, Yantosca RM, Sulprizio MP, Thompson AM, Wennberg PO, Crounse JD, St. Clair JM, Cohen RC, Laughner JL, Dibb JE, Hall SR, Ullmann K, Wolfe GM, Pollack IB, Peischl J, Neuman JA, Zhou X (2016) Why do models overestimate surface ozone in the Southeast United States? Atmos Chem Phys 16(21):13,561–13,577. doi:10.5194/acp-16-13561-2016

U.S. EPA (2014) Health Risk and Exposure Assessment for Ozone First External Review Draft. Technical report, Washing- ton, DC, https://www3.epa.gov/ttn/naaqs/standards/ozone/data/ 20140829healthrea.pdf

U.S. EPA (2015a) National Ambient Air Quality Standards for Ozone Final Rule

U.S. EPA (2015b) Overview of EPA's Updates to the Air Quality Standards for Ground-Level Ozone. Technical report, https://www. epa.gov/sites/production/files/2015-10/documents/overview of 2015 rule.pdf

USEPA (1999a) Regulatory Impact Analysis - Control of Air Pollution from New Motor Vehicles : Tier 2 Motor Vehicle Emissions Standards and Gasoline Sulfur Control Requirements. Tech. rep., United States Environmental Protection Agency

USEPA (1999b) Technical Support Document for the Tier 2 / Gasoline Sulfur Ozone Modeling Analyses. Technical Report December, http://www.epa.gov/scram001/reports/tsd for tier 2 gasoline sulfur rule.pdf

USEPA (2000a) Regulatory Impact Analysis: Heavy-Duty Engine and Vehicle Standards and Highway Diesel Fuel Sulfur Control Requirements, Chapter 3. Technical report, United States EPA

USEPA (2000b) Technical Support Document for the Heavy-Duty Engine and Vehicle Standards and Highway Diesel Fuel Sulfur Control Requirements: Air Quality Modeling Analyses. Technical report, United States EPA

USEPA (2005) Technical Support Document for the Final Clean Air Interstate Rule Air Quality Modeling. Tech. Rep. March, United States Environmental Protection Agency

USEPA (2009) NOx Budget Trading Program - Basic Information. Tech. rep., http://www.epa.gov/airmarkets/progsregs/nox/ docs/NBPbasicinfo.pdf

USEPA (2012a) Acid Rain Program. 40 Code of Federal Regulations Parts 72-78

USEPA (2012b) Motor vehicle emissions simulator (moves) user guide for moves2010b. http://www.epa.gov/otaq/models/moves/

USEPA (2012c) Motor vehicle emissions simulator (moves) v. 2010b. http://www.epa.gov/otaq/models/moves/

USEPA (2013) United states epa air markets program data. http://ampd.epa.gov/ampd/

van Erp AM, O'Keefe R, Cohen AJ, Warren J (2008) Evaluating the effectiveness of air quality interventions. J Toxicol Environ Health Part A 71(9-10):583–7. doi:10.1080/15287390801997708, http://www.ncbi.nlm.nih.gov/pubmed/18569630

Vijayaraghavan K, DenBleyker A, Ma L, Lindhjem C, Yarwood G (2014) Trends in on-road vehicle emissions and ambient air quality in Atlanta, Georgia, USA, from the late 1990s through 2009. J Air Waste Manag Assoc 64(7):808–816. doi:10.1080/10962247.2014.892039

Weber RJ, Guo H, Russell AG, Nenes A (2016) High aerosol acidity despite declining atmospheric sulfate concentrations over the past 15 years. Nat Geosci 9(April):1–5. doi:10.1038/NGEO2665

Xie Y, Elleman R, Jobson T, Lamb B (2011) Evaluation of O 3 -NO x - VOC sensitivities predicted with the CMAQ photochemical model using Pacific Northwest 2001 field observations. J Geophys Res 116(D20):D20,303. doi:10.1029/2011JD015801

Xing J, Zhang Y, Wang S, Liu X, Cheng S, Zhang Q, Chen Y, Streets DG, Jang C, Hao J, Wang W (2011) Modeling study on the air quality impacts from emission reductions and atypical meteorological conditions during the 2008 Beijing Olympics. Atmos Environ 45(10):1786–1798. doi:10.1016/j.atmosenv.2011.01.025, http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=59334233& site=ehost-live&scope=cite$\delimiter"026E30F$nhttp://www.sciencedirect.com/science/article/pii/S135223101100032X

Yang Y, Wilkinson J, Russell A (1997) Fast, direct sensitivity analysis of multidimensional photochemical models. Environ Sci Technol 31:2965–2976

Zhang W, Trail MA, Hu Y, Nenes A, Russell AG (2015) Use of high- order sensitivity analysis and reduced-form modeling to quantify uncertainty in particulate matter simulations in the presence of uncertain emissions rates: A case study in Houston. Atmos Environ 122:103–113. doi:10.1016/j.atmosenv.2015.08.091

Zhou W, Cohan DS, Napelenok SL (2013) Reconciling NOx emissions reductions and ozone trends in the U.S., 20022006. Atmos Environ 70(x):236–244. doi:10.1016/j.atmosenv.2012.12. 038, http://linkinghub.elsevier.com/retrieve/pii/S1352231013000 058

CHAPTER 6. AIR QUALITY MODELING FOR ACCOUNTABILITY RESEARCH: OPERATIONAL, DYNAMIC, AND DIAGNOSTIC EVALUATION[5]

6.1 Abstract

Photochemical grid models play a central role in air quality regulatory frameworks, including in air pollution accountability research, which seeks to demonstrate the extent to which regulations causally impacted emissions, air quality, and public health. There is a need, however, to develop and demonstrate appropriate practices for model application and evaluation in an accountability framework. We employ a combination of traditional and novel evaluation techniques to assess four years (2001-02, 2011-12) of simulated pollutant concentrations across a decade of major emissions reductions using the Community Multiscale Air Quality (CMAQ) model. We have grouped our assessments in three categories: *Operational* evaluation investigates how well CMAQ captures absolute concentrations; *dynamic* evaluation investigates how well CMAQ captures changes in concentrations across the decade of changing emissions; *diagnostic* evaluation investigates how CMAQ attributes variability in concentrations and sensitivities to emissions between meteorology and emissions, and how well this attribution compares to empirical statistical models. In this application, CMAQ captures O_3 and $PM_{2.5}$ concentrations and change over the decade in the Eastern United States similarly to past CMAQ applications and in line with model evaluation guidance; however, some $PM_{2.5}$ species—EC, OC, and sulfate in particular—exhibit high biases in various months. CMAQ-simulated $PM_{2.5}$ has a high bias in winter months and low bias in the summer, mainly due to a high bias in OC during the cold months and low bias in OC and sulfate during the summer. Simulated O_3 and $PM_{2.5}$ changes across the decade have normalized mean bias of less than 2.5% and 17%, respectively. Detailed comparisons suggest biased EC emissions, negative wintertime

[5] A modified version of this chapter is under review for publication as: Henneman, L.R.F., Liu, C., Hu, Y., Mulholland, J.A., Russell, A.G. "Air quality modeling for accountability research: Operational, dynamic, and diagnostic evaluation." The supplementary material is included in Appendix A.

SO_4^{2-} sensitivities to mobile source emissions, and incomplete capture of OC chemistry in the summer and winter. Photochemical grid model-simulated O_3 and $PM_{2.5}$ responses to emissions and meteorology across the decade match results from receptor-based, statistical regression models. $PM_{2.5}$ sensitivities to mobile source emissions in the summertime have decreased substantially, but wintertime mobile sensitivities remain largely unchanged because decreases in negative SO_4^{2-} sensitivities match decreases in positive sensitivities from other constituents. Similarly, NO_X emissions have led to decreased summertime O_3 and increased wintertime O_3 because of opposite sensitivities. Overall, results show that emissions reductions improved air quality across the domain and remain a viable option for improving future air quality.

6.2 Introduction

Rigorous assessment of current regulations plays an important role in shaping future policy decisions. To this end, air pollution accountability research seeks to causally attribute regulatory interventions to changes in emissions, air quality, exposure/dose, and public health—links in the so-called *Accountability Chain*—to specific regulations (Health Effects Institute 2003). Causal linkages are difficult to establish, however, due to uncertainties in the links between each step. For example, the link between emissions and air quality is uncertain due to simultaneous variability in meteorology and other emissions sources. Photochemical grid models (PGMs) offer a mechanistic approach for quantifying these relationships for accountability research.

PGMs play a vital role in air quality regulatory frameworks in the United States and Europe (National Research Council 2004), and are increasingly used elsewhere (Hu et al. 2016; Liu et al. 2010; Xing et al. 2015). These models, alternatively called chemical transport models (CTMs), simulate emissions, transport, formation, and fate of multiple air pollutants in the atmosphere. Results from PGMs are applied in a variety of capacities, including regulatory applications (Simon et al. 2013), air quality forecasting (Odman et al. 2007), atmospheric chemistry research (Brune et al. 2016; Park et al. 2004; Travis et al. 2016), and exposure studies (Bravo et al. 2012; Di et al. 2016; Fann et al. 2012; Muller et al. 2009; U.S. EPA 2015b).

Performance evaluation provides information relating to the reliability of model results, the magnitude of biases and uncertainties, and the applicability of model results to

various problems. Existing literature contains suggested model performance metrics and benchmarks. U.S. EPA (1991), for example, recommended a set of performance criteria for 1-hour ozone. Boylan and Russell (2006) compiled results from a variety of studies using PGM modeling and considered how such models are used in the regulatory process to develop concentration-dependent performance guidance for particulate matter (PM) with diameter less than 2.5 μm ($PM_{2.5}$). Simon et al. (2012) later summarized 69 studies containing model evaluation information to provide a contextual background for ozone, PM and PM species modeling performance expectations. Emery et al. (2016) used results from Simon et al. (2014) and more recent studies to recommend numerical criteria and goals for ozone, $PM_{2.5}$, and its species, and further suggested temporal and spatial scales in which these recommended numbers apply, i.e., about 1000 km, less than one month for ozone, and one month or one season for $PM_{2.5}$ and $PM_{2.5}$ species.

Dennis et al. (2010) proposed a four-tier framework—which is reflected in regulatory language currently under development by the US EPA (U.S. EPA 2014)—for model performance evaluation: operational, dynamic, diagnostic, and probabilistic. Among the four tiers, operational evaluation, in which modeled results are compared against corresponding measured data, is applied most frequently. Dynamic evaluation results, which quantify how well models capture the impact of emissions changes, have important regulatory implications because PGMs are often applied to estimate expected changes in concentration due to proposed controls (e.g., U.S. EPA 1998, 2005). Diagnostic evaluation assesses processes and components of the modeling system, and probabilistic evaluation compares modeled and observed distributions of specific variables, instead of focusing on matched spatio-temporal data. A detailed knowledge of model limitations in this domain is critical to regulatory decision-making as manifested in control strategy development.

Unlike operational and dynamic evaluations, diagnostic evaluations often rely on comparisons with more empirical methods and quantities inferred or derived from observations (e.g., Digar et al. 2013; Xie et al. 2011). Godowitch et al. (2011), for example, compared ozone production efficiency and results from a metric to assess wind profiles to evaluate CMAQ's ability to simulate these variables. Marmur et al. (2006 and 2009) compared various observation and CMAQ-based source apportionment techniques to assess their agreement. While neither the empirical or PGM-based methods are infallible,

these comparisons provide important points of context for unobservable quantities, such as sensitivities to emissions and meteorology.

Notably, the four-tier framework has been applied in the Air Quality Model Evaluation International Initiative (AQMEII—(Hogrefe et al. 2015; Rao et al. 2011; Xing et al. 2015)). Im et al. (2015a, 2015b), for example, evaluated multiple models' abilities to simulate O_3 and $PM_{2.5}$ over United States and Europe. Work under this initiative has applied the multi-tiered framework for model evaluation, with an emphasis on diagnostic and probabilistic evaluations (Solazzo et al. 2017b, 2017a). (Solazzo and Galmarini 2016) separated model error into components and related these to model processes at various time scales, and determined that long-term processes and input fields (e.g., emissions and boundary conditions) contributed most to the model error.

Modeling requirements for accountability studies align with needs for detailed model assessments over periods of changing emissions and meteorology. Several studies have applied dynamic evaluation to simultaneously test impacts of emissions reductions on air quality and assess the model's ability to capture the reduction (Banzhaf et al. 2015; Daskalakis et al. 2016; Foley et al. 2015a; Gégo et al. 2008; Godowitch et al. 2010; Simon et al. 2014), and a few have applied diagnostic evaluation techniques to attribute variability in concentrations between meteorology and emissions differences (Foley et al. 2015b; Gilliland et al. 2008; Godowitch et al. 2008; Napelenok et al. 2011). Most studies to date in this domain have focused on ozone concentrations; few studies have assessed $PM_{2.5}$ and $PM_{2.5}$ species.

As part of an accountability study on the effectiveness of regulations implemented under the 1990 Clean Air Act Amendments, we applied the Community Multiscale Air Quality model (CMAQ) version 5.0.2 over a domain covering the eastern United States to 2 two-year periods spanning a period of large emissions reductions: 2001-2002 and 2011-2012. With the Decoupled Direct Method (DDM) extension, we calculated air pollution concentration sensitivities to electricity generating unit (EGU) and on-road mobile sources. Results from CMAQ were evaluated using multiple evaluation tiers. First, we apply an operational evaluation to answer the question: *How well does the PGM, as applied, capture observed concentrations?* Next, we apply a dynamic evaluation to answer the question: *Does the PGM capture observed air quality changes?* In a diagnostic analysis, we ask two

novel questions: *Can the PGM reproduce empirically-derived pollutant variability attributed to emissions and meteorology?* and *How well does the PGM reproduce the air pollutant sensitivities to emissions as simulated via empirical statistical modeling?* Together, the analysis finds that observed air quality changes during the period are attributable primarily to emissions reductions, and supports the use of the PGM in an accountability applications.

6.3 Methods

6.3.1 CMAQ modeling and inputs

CMAQ (version 5.0.2, using CB05; (Byun and Schere 2006, https://www.cmascenter.org/index.cfm)) simulations were performed for the Eastern continental US (Figure A-1). The modeling domain was chosen to focus on air quality in the southeastern US, where detailed measurements are available from the Southeastern Aerosol Characterization (SEARCH) Study (Blanchard et al. 2013) and empirical accountability analysis has previously been conducted for comparison (Henneman et al. 2015, 2017). Simulations were performed over the Eastern US for 2001-2002 and 2011-2012, using a 12 km x 12 km resolution grid with 13 vertical layers of different thicknesses (thinner near the ground). We used the EPA SMOKE platform (https://www.epa.gov/air-emissions-modeling/2011-version-6-air-emissions-modeling-platforms) to prepare emissions from the 2002 and 2011 National Emission Inventories (NEI) for 2001-2002 and 2011-2012 modeling years, respectively. We applied the Weather Research Forecast (WRF) model version 3.6.1 to generate meteorological fields. The EPA SMOKE platform included biogenic and dust emissions. A fixed profile was used for boundary conditions.

6.3.2 Observations

Domain-wide observational data for 2001-2002 and 2011-2012 were downloaded from the US EPA Air Quality System (AQS, https://www.epa.gov/ttn/airs/airsaqs/detaildata/downloadaqsdata.htm) and used to evaluate how well the model simulated ozone and $PM_{2.5}$. $PM_{2.5}$ species (sulfate (SO_4^{2-}), nitrate (NO_3^-), ammonium (NH_4^+), organic carbon (OC), and elemental carbon (EC)) observations were taken from EPA's Chemical Speciation Network (CSN). In a detailed analysis of model performance in the Southeast, we used observations of the same species

from the SEARCH network's four urban-rural pairs (8 monitoring locations) (Blanchard et al. 2013). Monitoring sites used in the analysis are plotted in Figure A-1.

6.3.3 Operational evaluation

Following recommendations by Emery et al. (2016) and others, we calculated normalized mean bias (NMB), normalized means error (NME), and correlation coefficient (r). NMB, NME, and correlation (*r*) are defined as:

$$\mathrm{NMB} = \frac{\sum(P_i - O_i)}{\sum O_i} \quad \text{(Eq. 6.1)}$$

$$NME = \frac{\sum|\mathrm{P_i} - O_i|}{\sum O_i} \quad \text{(Eq. 6.2)}$$

$$R = \frac{\sum[(\mathrm{P_i} - \overline{\mathrm{P}})\times(O_i - \bar{O})]}{\sqrt{\sum(P_i - \bar{P})^2 \times \sum(O_i - \bar{O})^2}} \quad \text{(Eq. 6.3)}$$

O_i is the *i'th* observation used in the evaluation, P_i is the associated model simulated value, and $\bar{P}$ and $\bar{O}$ are the mean values of model simulation and observation, respectively. No cutoff was applied for evaluation of daily pollutant metrics in the present study, and we applied a cutoff of 40 ppb for hourly ozone. R is preferred over the alternative, R^2, because the sign shows directionality.

NMB provides an average relative magnitude of model bias and its direction. An NMB with a small absolute value, regardless of its sign, indicates that the model can capture averaged values over spatial and/or temporal scales for which NMB is calculated. In addition, because it is normalized by observations, NMB evaluates in a relative—as opposed to absolute—scale, and facilitates inter-study and interspecies comparison. Intra-study model evaluation on an absolute scale is of interest, for example, in an evaluation of how biases in various $PM_{2.5}$ species contribute to total $PM_{2.5}$ mass. Mean bias (MB) can provide such information:

$$MB = \frac{\sum(\mathrm{P_i} - \mathrm{O}_i)}{N} \quad \text{(Eq. 6.4)}$$

where *N* is the number of simulation-observation pairs.

6.3.4 Dynamic evaluation

Dynamic evaluation was conducted using the same monitoring data as the operational evaluation. We calculated observed and modeled differences in daily concentrations across each two-decade period (i.e., 2001-2011 and 2002-2012) by pairing

by day of year. While O_3 and $PM_{2.5}$ were measured every day throughout 2001-2012 at AQS sites, PM species were sampled every three days after 2007 at CSN sites. Because this led to sparse data sets of daily pairs, we averaged available observation and modeled concentrations by month at each site to facilitate analysis for dynamic evaluation of $PM_{2.5}$ species. The use of monthly averages means fewer points were available for comparison relative to O_3 and $PM_{2.5}$, but the tradeoff is necessary due to sparse observation datasets.

The change of concentration could be positive and negative, whereas the measured concentration is always positive. Therefore, we modified the operational definitions of NME and NMB slightly by using differences instead of absolute concentrations and taking absolute value in the denominator, and changed all metrics to reflect use of the change in concentrations that is,

$$NMB_{\Delta} = \frac{\sum(\Delta P_i - \Delta O_i)}{\sum|\Delta O_i|} \quad \text{(Eq. 6.5)}$$

$$NME_{\Delta} = \frac{\sum|\Delta P_i - \Delta O_i|}{\sum|\Delta O_i|} \quad \text{(Eq. 6.6)}$$

$$MB_{\Delta} = \frac{\sum(\Delta P_i - \Delta O_i)}{N} \quad \text{(Eq. 6.7)}$$

$$r_{\Delta} = \frac{\sum[(\Delta P_i - \Delta\bar{P})\times(\Delta O_i - \Delta\bar{O})]}{\sqrt{\sum(\Delta P_i - \overline{\Delta P})^2 \times \sum(\Delta O_i - \Delta\bar{O})^2}} \quad \text{(Eq. 6.8)}$$

where the subscript Δ denotes the *dynamic* version of each metric.

6.3.5 Diagnostic evaluation

The diagnostic evaluation portion of this work has two goals: 1) to quantify and separate the contributions of meteorological and emissions variability across the decade to CMAQ-DDM modeled concentrations and sensitivities and 2) to compare these contributions to similar estimates using observed data and statistical modeling. Since PGMs and statistical models cannot be evaluated against observations to confirm or refute the findings of emissions or meteorological influences, this evaluation serves to establish a confidence level related to how well two different approaches produce similar results.

6.3.5.1 CMAQ cross-simulations and sensitivities

In observed ambient data, it is difficult to separate the impacts of emissions and meteorology—two primary air quality drivers. Previous studies have used statistical detrending methods to estimate how differences in inter-annual meteorological conditions

affect O_3 concentrations (Blanchard et al. 2014; Camalier et al. 2007; Cox and Chu 1993; Henneman et al. 2015; Kuebler et al. 2001; Rao et al. 1997). In PGMs, researchers have employed cross simulations—i.e., re-running the model in different years and substituting alternate years' emissions or meteorology—to separate emissions and meteorology effects on changing concentration (Foley et al. 2015b; Godowitch et al. 2008).

In addition to the four base years, we conducted a year-long simulation with switched meteorology and anthropogenic emissions; i.e., we ran a year-long simulation driven by 2001 anthropogenic emissions and 2011 meteorology and biogenic emissions. We estimated the impacts of meteorological and emissions variability on CMAQ outputs by subtracting results from between the modeling runs (Table 6-1). Each CMAQ run is defined by its meteorology and emissions years, e.g., 01M:11E denotes the run with 2001 meteorology and 2011 emissions. Differences designate the effect being investigated and the year and input held constant, e.g., E.2001m represents the impact of changing emissions (E) using 2001 meteorology (m).

Table 6-1. Definitions of scenarios used for isolating emissions and meteorological effects on CMAQ output. Statistical and CMAQ columns show whether the scenario was included in statistical and CMAQ modeling. STM is the short-term meteorology from the detrending model.

Difference	Abbreviation	Description	CMAQ	Statistical
01M:01E - 11M:11E	Total	Total change	Yes	Yes
01M:01E - 01M:11E	E.2001m	Emissions effect (2001 meteorology)	Yes	Yes
01M:11E - 11M:11E	M.2011e	Meteorology effect (2011 emissions)	Yes	Yes
	STM	Short-term meteorology	No	Yes

Modeled concentration sensitivities to emissions are calculated using the direct-decoupled method (DDM) in CMAQ (Hakami et al. 2004). CMAQ-DDM employs a Taylor Series expansion to calculate semi-normalized first-order sensitivities ($S_{i,j}$) to emissions from source *j* for gaseous and particulate species *i* (Cohan et al. 2005; Napelenok et al. 2006). $S_{i,j}$ is effectively determined as (Yang et al. 1997):

$$S_{i,j} = \frac{\partial C_i}{\partial \alpha_j} \quad \text{(Eq. 6.9)}$$

Here, α_j, which has a nominal value of 1, represents the relative fraction of emissions from source *j* in the base calculation. $S_{i,j}$ are in the same units of concentration, and provide an estimate of pollutant response to precursor emissions assuming a linear

system. Previous studies have shown that the linearity assumption holds for domain-wide emissions changes up to 30-50% depending on species (Cohan et al. 2005; Hakami et al. 2004).

DDM sensitivities have been used in research (Cohan 2004; Henneman et al. 2017; Souri et al. 2016) and regulatory (Simon et al. 2013; U.S. EPA 2015a) applications as both a method to attribute portions of modeled concentrations to specific sources and to assess the impacts of potential controls. Here, we show sensitivities for both the base case and how they are affected by meteorology and emissions changes.

6.3.5.2 Statistical modeling

We employ results from two varieties of empirical statistical models for diagnostic comparisons previously developed and applied for accountability analyses (Henneman et al. 2015, 2017). We employ the term 'empirical' here to emphasize that the statistical models describes complex relationships using statistical correlations between observed and modeled quantities, and not through the mechanistically-based differential equations that drive deterministic air quality models (including CMAQ). The first, meteorological detrending, separates the influence of daily meteorological variability on air pollution concentrations from other factors (see Henneman et al. (2015) for details). Log daily ambient pollutant concentrations were split into five components with variability at different periods: long-term (LT, period > 1 year), seasonal (S, period = 1 year), weekly-holiday (WH, period = 7 days to 1 year), short-term meteorological (STM, period < 1 year), and white noise (WN, period = 1 day):

$$log\ [C(t)] = LT(t) + S(t) + WH(t) + STM(t) + WN(t) \qquad \text{(Eq. 6.10)}$$

We used multiple Kolmogorov-Zurbenko filters to separate and remove the LT and S portions of the original ambient concentration and meteorology signals with terms for weekdays and holidays. The remaining daily concentrations deviations were regressed against daily meteorology (temperature, relative humidity, wind speed, and rainfall) deviations and interactions with weekdays, holidays, and months. The fit of this regression without the weekday-holiday terms and transformed to concentration units from exponential concentration units approximates the contribution of meteorological variability to observed ambient concentrations, STM. The detrending method was applied to all available data from the SEARCH network.

The second statistical approach (described and evaluated in detail in Henneman et al., 2017) estimates source and species-specific emissions contributions to observed air pollution concentrations at JST. To this end, we used pollutant-specific multivariate models of the form:

$$C = ß_0 + ß_E(\boldsymbol{E}) + ß_M(\boldsymbol{M}) + ß_{EM}(\boldsymbol{EM}) + \in \quad \text{(Eq. 6.11)}$$

where C is the vector of daily ambient concentration, $\boldsymbol{E}$ is the matrix of daily mobile and EGU emissions, $\boldsymbol{M}$ is the matrix of daily observed meteorological parameters at JST (and supplemented with data from nearby Hartsfield-Jackson International Airport), $\boldsymbol{EM}$ represents interactions between emissions and meteorology, and $\in$ is the vector of model residuals. We use the model to separate the influences of emissions of multiple species from mobile sources and local and regional EGUs. Atlanta-area (the 20-county $PM_{2.5}$ non-attainment area) mobile emissions were modeled using MOVES2010b with inputs provided by the Georgia Environmental Protection Division. EGU emissions for the Southeast (Alabama, Georgia, Mississippi, North Carolina, South Carolina, and Tennessee) were downloaded from EPA's Air Markets Program Database (U.S. EPA 2016) and split into two groups: plants inside the non-attainment area and plants outside. We extended the analysis from O_3 and $PM_{2.5}$ investigated in Henneman et al. (2017) to multiple gaseous (NO_2, SO_2, and CO) and $PM_{2.5}$ (SO_4^{2-}, NH_4, NO_3^-, EC, and OC) species. This method captures the impacts of local and regional emissions and meteorology, but not impacts of emissions transported into the region.

We again applied cross-simulations to estimate impacts of emissions changes on concentrations in each species' statistical model (Eq. 6.11); model parameters (ß's) and meteorological variables were held constant and emissions were switched to achieve combinations in Table 6-1. The difference in meteorological influence on the two years was estimated as the difference between 2001 and 2011 STM using the detrending model. We derived pollutant sensitivities to source-specific emissions (comparable to sensitivities estimated by CMAQ-DDM) by summing model parameter-covariate combinations in Eq. 6.11 that included the emissions variable (β_E, β_{EM}). This yielded daily sensitivities for each pollutant in the same units as ambient concentration.

6.4 Results

We split the evaluation results into three parts: operational, dynamic, and diagnostic. The first two focus solely on CMAQ and ambient observations, and the third investigates how much change in concentrations and sensitivities CMAQ attributes to meteorology vs. emissions, and how these results compare to the statistical model results.

Emery et al. (2016), using terminology advanced by Boylan and Russell (2006), suggest benchmarks at two stringency levels—*criteria* and *goals*—for NMB, NME, and *r*. While any modeling effort should strive for low bias, low error, and high correlation, the authors define these metric benchmarks to establish references across modeling efforts employing varying methodologies. Their purpose is not to serve as strict pass/fail tests for model evaluation, but to help in understanding model uncertainties in a structured setting. The metrics were set based on results from 31 published modeling efforts; *criteria* and *goal* values were achieved by 33% and 67% of the included studies, respectively. We discuss our domain-wide results in relation to these benchmarks to help establish the quality of the modeling used here compared to previous studies.

6.4.1 Operational Evaluation

Annual domain-wide operational evaluation of CMAQ model performance, yielded similar results across the four years for each pollutant (Table 6-2, Table A-1 and Table A-2 for SEARCH sites and JST, respectively).

Table 6-2. Domain-wide CMAQ operational performance evaluation for ozone (hourly and MDA8h), $PM_{2.5}$ (24-hr), and PM species (24-hr) for 2001, 2002, 2011, and 2012. Ozone and $PM_{2.5}$ are evaluated against all AQS sites, and $PM_{2.5}$ species are evaluated against CSN sites. N denotes the number of monitor-day pairs in each comparison (N values reflect the dataset before the cutoff is applied). Hourly ozone is evaluated with a cutoff of 40 ppb except for r.

Performance Metric/Year	Ozone (hourly)	Ozone MDA8h	$PM_{2.5}$	EC	OC	SO_4	NO_3	NH_4
2001								
N	1373143	189658	109339	3097	3097	7120	6630	4424
NMB	-13%*	5.2%†	0.84%†	124%	-12%†	-24%*	15%†	-3%†
NME	23%*	20%*	48%*	135%	57%*	48%*	74%*	58%
r	0.66*	0.68*	0.44*	0.47‡	0.26‡	0.51*	0.67‡	0.50*
2002								
N	1369293	193224	111348	8347	8347	13888	13174	10803
NMB	-13%*	4.9%†	-4.4%†	113%	-20%*	-28%*	-4%†	-18%*
NME	23%*	21%*	48%*	125%	63%*	45%*	61%*	45%*
r	0.67*	0.72*	0.38	0.52‡	0.17‡	0.57*	0.73‡	0.55*
2011								
N	1402937	197671	86036	9602	9602	9283	9521	9523
NMB	-13%*	4.3%†	-2.3%†	61%	7%†	-20%*	11%†	-1%†
NME	21%*	18%*	48%*	81%	61%*	42%*	61%*	50%*
r	0.63*	0.67*	0.40*	0.48‡	0.25‡	0.52*	0.73‡	0.61*
2012								
N	1185399	114445	71715	9669	9669	8920	9190	9194
NMB	-7.5%*	1.6%†	8.3%†	64%	12%†	-5%†	16%*	22%*
NME	11%†	16%†	48%*	85%	64%*	39%*	60%*	52%
R	0.63*	0.70*	0.48*	0.45‡	0.26‡	0.50*	0.78‡	0.68*

* Meets the criteria benchmark
† Meets the goal benchmark
‡ No benchmark available

6.4.1.1 Ozone

Ozone metrics (Table 6-2) at 1-hour and MDA8h time scales meet the NMB and NME criteria benchmarks (±15%, 25%, respectively) defined by Emery et al. (2016), but the correlation for hourly ozone is low. MDA8h ozone meets the stricter goal benchmarks (±5%, 15%, and 0.7 for NMB, NME, and r, respectively) for some of the metrics, and criteria benchmarks for all metrics. Evaluation benchmarks for MDA8h ozone were consistently better than for hourly ozone.

6.4.1.2 $PM_{2.5}$ and $PM_{2.5}$ species

Most of the $PM_{2.5}$ metrics across the four years meet the criteria (+/- 30%, 50%, and 0.4) and goal (+/- 10%, 35%, and 0.7) benchmarks, apart from correlation in 2002 and 2012 (Table 6-2). Simulated $PM_{2.5}$ is biased high from October through March and biased low from April through September in both 2001 and 2011 (Figure 6-1), consistent with conclusions made in previous studies (Foley et al. 2010; Park et al. 2006; Simon et al.

2012). Simulated concentrations in most months meet the relevant recommended criteria levels for NMB, NME, and *r*.

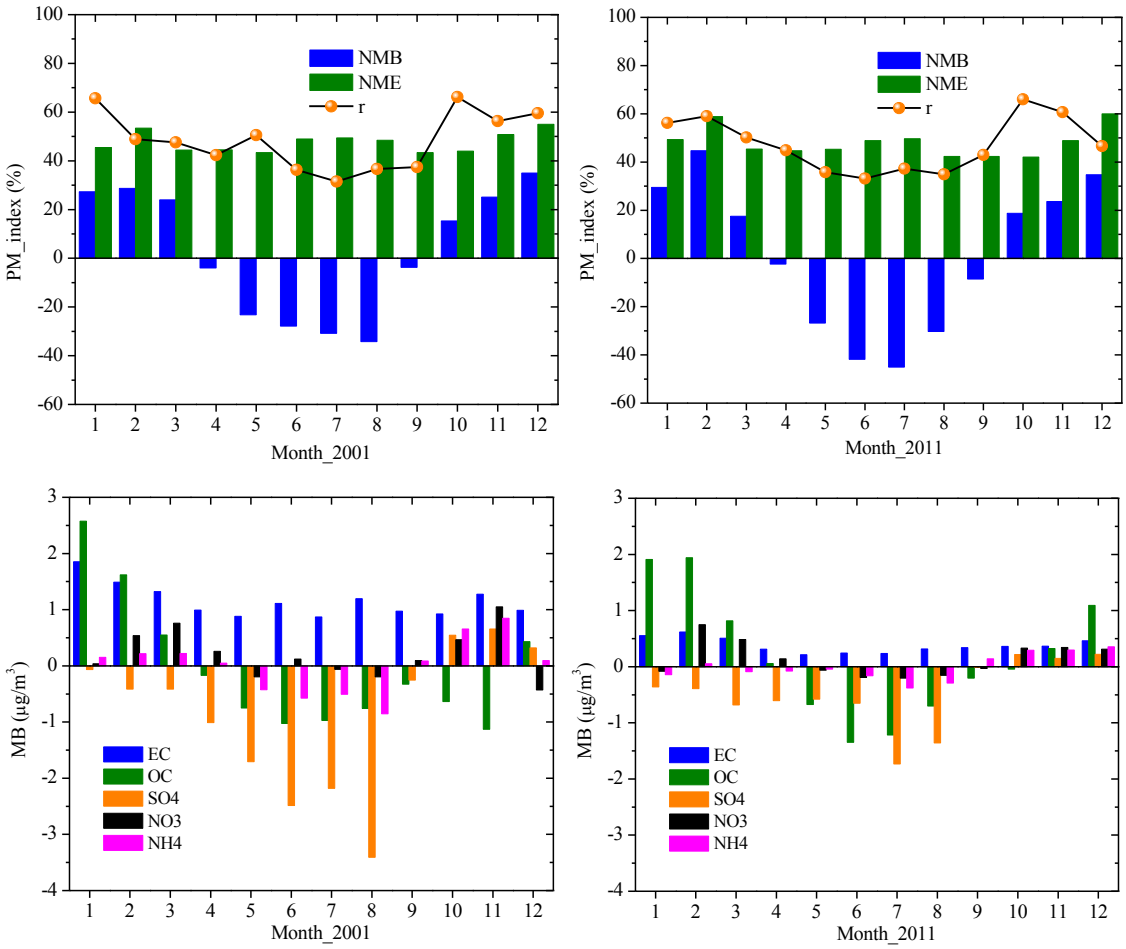


Figure 6-1. Monthly operational evaluation results for $PM_{2.5}$ mass. Left: 2001, right: 2011 (top). Monthly MB for $PM_{2.5}$ species. Left: 2001, right: 2011 (bottom).

Most $PM_{2.5}$ species performance statistics meet the criteria and goal benchmarks, which vary between species. EC, however, stands out; its NMB and NME exceed the suggested criteria (±40 and 75%, respectively), though results improve in 2011-12 compared to 2001-02. OC NMB and NME meet the criteria for all years, but OC consistently has a low R value.

$PM_{2.5}$ species MB decreases across most species and months from 2001 to 2011 (Figure 6-1). In 2011, OC dominates bias in January (1.9 μg/m^3), February (1.9 μg/m^3), and December (1.1 μg/m^3). OC bias remains high in May (-0.67 μg/m^3), June (-1.3 μg/m^3), July (-1.2 μg/m^3), and August (-0.70 μg/m^3) in 2011, and sulfate bias contributes comparatively to the total in these months (-0.58 μg/m^3, -0.65 μg/m^3, -1.7 μg/m^3, and -1.4

$\mu g/m^3$, respectively). MB of other species is generally low; EC, OC, nitrate, and ammonium contribute to the cold month overestimation, and OC, sulfate, and ammonium contribute to the warm month underestimation (Figures S2 to S3, and further discussion in the supplemental).

6.4.2 Dynamic evaluation

Ozone NMB_Δ and r_Δ meet the goal and criteria benchmarks, respectively (Table 6-3, Table A-3 and Figure A-7 for JST), though the benchmarks were not originally developed for assessing differences (Δ's) between model simulations between years. NME_Δ, however, is higher than the benchmarks for all species.

Table 6-3. Domain-wide CMAQ dynamic performance evaluation for MDA8h ozone, $PM_{2.5}$, and PM species for 2001-2011 and 2002-2012. All AQS sites are used for ozone and $PM_{2.5}$, and CSN sites are used for the species performance analyses. N denotes the number of sites in each comparison. Species differences are taken as the monthly site averages due to 1-in-3 sampling that left few days across the decade the same.

Performance Metric/Year	Ozone MDA8h	$PM_{2.5}$	EC	OC	SO_4	NO_3	NH_4
2001-2011							
N_Δ	134451	22747	311	311	365	359	358
NMB_Δ	-2.43%†	-0.27%†	-204%	63%	18%*	-3.7%†	6.6%†
NME_Δ	79.5%	90.6%	224%	101%	54%	94%	64%
R_Δ	0.60*	0.53*	0.42‡	0.34‡	0.59	0.51‡	0.41
2002-2012							
N_Δ	80934	18922	799	799	883	875	884
NMB_Δ	-1.02%†	16.7%*	-128%	70%	43%	30%*	45%
NME_Δ	80.8%	90.7%	173%	104%	54%	90%	58%
R_Δ	0.60*	0.53*	0.40‡	0.32‡	0.68	0.49‡	0.53

* Meets the criteria benchmark
† Meets the goal benchmark
‡ No benchmark available

Correlation results meet the criteria values for $PM_{2.5}$ and all species except OC (Table 6-3). In the 01:11 comparison, all species meet the criteria values for NMB_D except EC and OC. Between the two comparisons (01:11 and 02:12), the 01:11 simulation better captures observed SO_4, NO_3, and NH_4. EC and OC results are consistent between 01:11 and 02:12. This difference is reflected in monthly dynamic MB_D results (Figure 6-2), which show isolated peaks for SO_4, NO_3, and NH_4 in 01•11. In the 02•12 comparison, however, each of these species exhibits high MB_D for multi-month periods throughout the year.

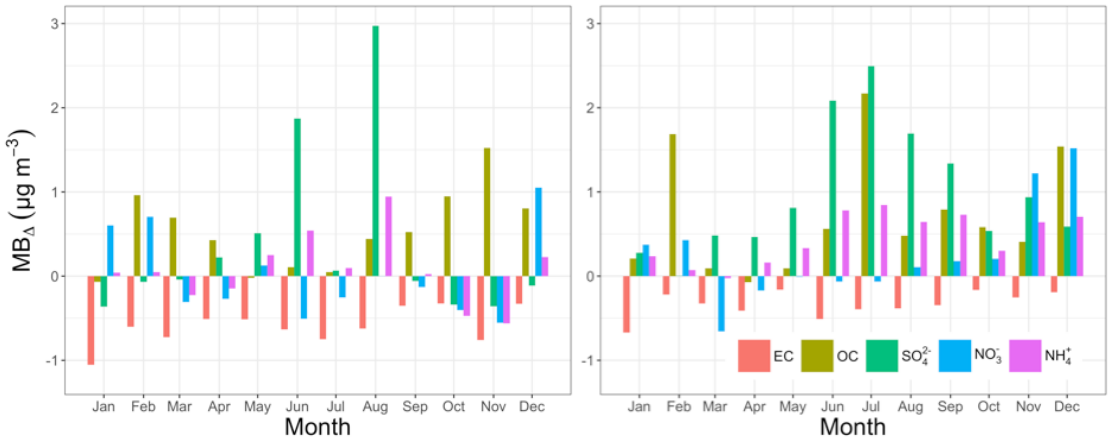


Figure 6-2. Results for the dynamic evaluation of $PM_{2.5}$ species between 2001 and 2011 (left) and 2002 and 2012 (right), using monthly-averaged observations across all CSN monitors.

6.4.3 Diagnostic evaluation

6.4.3.1 Concentration changes attributable to emissions and meteorology

We compared impacts of emissions and meteorological variability on modeled outputs using cumulative distribution functions (CDFs) for each of the scenarios (Table 6-2) at all SEARCH sites (Figure 6-3). Daily differences have been normalized by 2011 concentrations to facilitate comparisons between CMAQ-modeled data and statistically modeled data. We used CMAQ-simulated concentrations to normalize CMAQ differences and observed concentrations to normalize statistically modeled differences. A difference of 10 ppb on a day with a 2011 O_3 concentration of 50 ppb would yield a normalized difference of 0.2. The comparison shows that the CMAQ and statistical model results generally agree across the years. Above the median ozone change between 2001 and 2011, meteorology (green line for CMAQ, red for statistical) is found to play a greater role in increasing concentrations than emissions (blue and orange). CMAQ estimates a slightly negative contribution of median emission changes to ozone concentrations, and the statistical model finds a median contribution of zero.

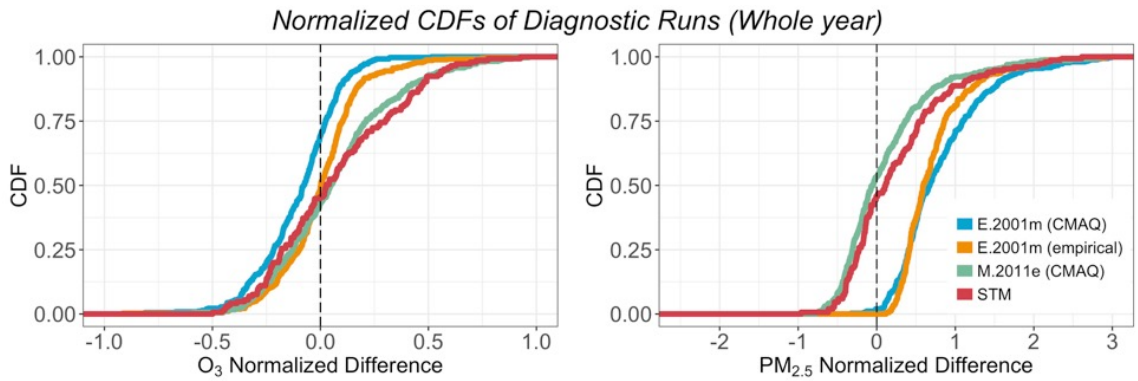


Figure 6-3. CDFs of differences in O_3 and $PM_{2.5}$ at JST captured by CMAQ and the statistical models. Differences have been normalized by 2011 concentrations. E.2001m (CMAQ) and E.2001m (statistical) are the impact of emissions on changes between 2001 and 2011 based on 2001 meteorology. Similarly, M.2011e (CMAQ) and STM (short-term meteorology derived using meteorological detrending) represent impacts on concentrations of meteorology changes.

At all SEARCH sites (Figure 6-4), the median impact of emissions-driven changes in concentrations (E.2001m) in January (July) was -0.18 ppb ppb^{-1} (0.12 ppb ppb^{-1}). January E.2001m exceeded meteorology-driven changes (M.2011e) across the distribution. In July, M.2011e exceeded E.2001m at the upper end of the distribution, meaning that emissions changes across the decade had less impact on O_3 concentrations at higher concentrations.

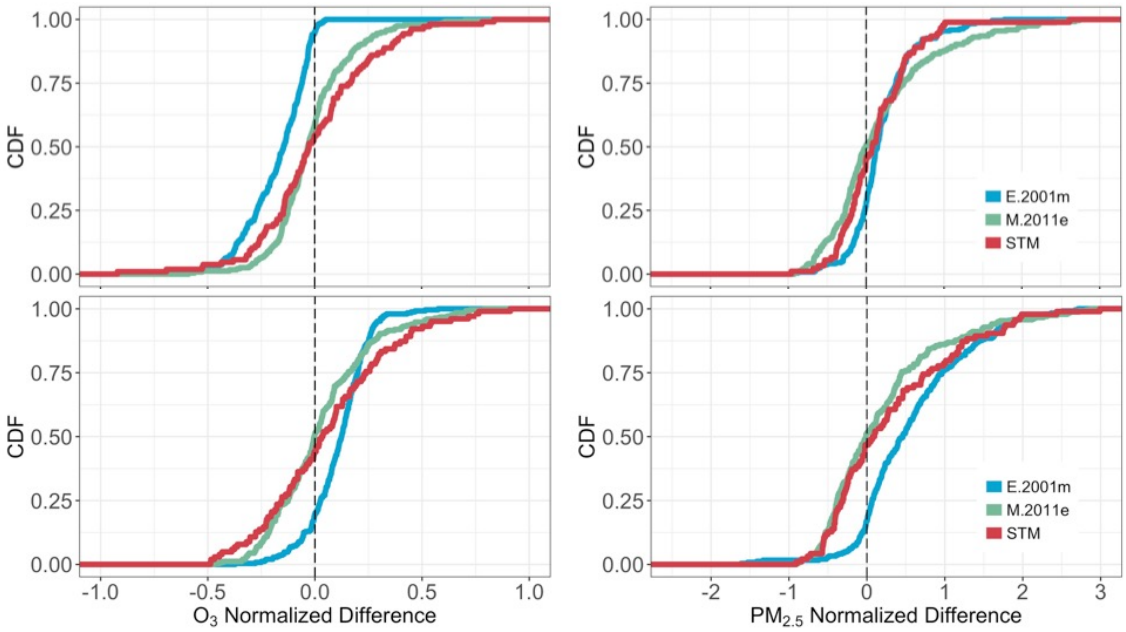


Figure 6-4. CDFs of differences in ozone and $PM_{2.5}$ concentrations between E.2001m, M.2011e, and STM at SEARCH sites in January (top) and July (bottom). Differences have been normalized by 2011 ozone concentrations.

Results for both January and July show general agreement between CMAQ-modeled (M.2011e) and statistically modeled (STM) meteorological impacts on concentration variability (Figure 6-4). In January, median meteorology impact is slightly negative, and in July, meteorological impact is slightly positive. STM results suggest a slightly larger (in magnitude) effect across the distribution than M.2011e.

The median E.2001m for $PM_{2.5}$ is 0.15 ug m^{-3} [µg m^{-3}]$^{-1}$ in January, and 0.46 ug m^{-3} [µg m^{-3}]$^{-1}$ in July. In both months, M.2011e and STM suggest a negligible impact of median meteorological variability. At higher percentiles in the January distribution, M.2011m imparts nearly two times the impact on concentrations compared to E.2001m, though the STM results suggest that this may be exaggerated. In July, emissions and meteorological effects are comparable above the 90th percentile.

6.4.3.2 Sensitivity changes attributable to emissions and meteorology

CMAQ DDM sensitivity magnitudes of ozone and $PM_{2.5}$ to source emissions generally decreased over time, though spatial plots by season show the complexity of emissions-air quality relationships and how they change over time (Figure 6-5). This is due largely to the decrease in emissions. The sensitivities as calculated by CMAQ-DDM are a first-order estimate of how much the absolute, not relative, concentrations of the pollutants

would change if all the associated emissions were removed. Thus, if the emissions in 2011 are 50% of what they were in 2001, and the system is purely linear, the sensitivity would be halved. However, given the complex nonlinear chemistry, reductions in emissions of multiple pollutants at different rates across the decade, and the nonhomogeneous spatial emissions changes, an appropriate approach to normalizing the sensitivities is not apparent.

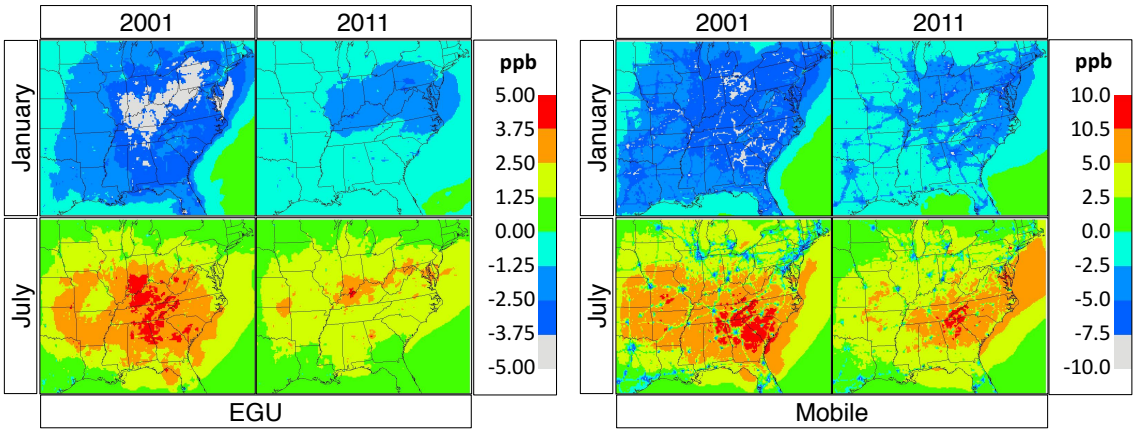


Figure 6-5. Monthly-averaged hourly ozone sensitivities to EGU (left) and on-road mobile emissions (right). Note that the scales are different.

Ozone sensitivities to EGU emissions tend to be negative in winter, positive in summer, and decrease in magnitude across the time of interest. Across the decade, for example, MDA8h O_3 sensitivities to EGU emissions in January increased from -0.8 ppb to -0.3 ppb, and sensitivities to on road mobile source emissions increased from -1.8 ppb to -0.8 ppb (Table 6-4). In July, MDA8h O_3 sensitivities to EGU emissions decreased from 3.4 ppb to 2.3 ppb, and on-road mobile sensitivities decreased from 7.7 ppb to 5.9 ppb.

Table 6-4. Domain-wide annual average CMAQ-DDM sensitivities in the base-case runs (01E:01M and 11E:11M).

	EGU		On-road mobile	
	2001	**2011**	**2001**	**2011**
O_3 (hourly—ppb)				
January	-2.0	-0.8	-3.9	-1.9
July	1.9	1.4	3.7	3.3
O_3 (MDA8h—ppb)				
January	-0.8	-0.3	-1.8	-0.8
July	3.4	2.3	7.7	5.9
$PM_{2.5}$ ($\mu g\ m^{-3}$)				
January	0.8	0.6	0.9	0.8
July	2.4	1.1	1.6	0.4

$PM_{2.5}$ sensitivities to EGU and on-road mobile sources decreased more in the summer than the winter (Table 6-4, Figure 6-6). Over the decade, winter on-road mobile sensitivities decreased from 0.9 µg m^{-3} to 0.8 µg m^{-3}, and summer sensitivities decreased from 1.6 µg m^{-3} to 0.4 µg m^{-3}, meaning that the dominant mobile contribution to $PM_{2.5}$ switched from summer to winter. Sulfate sensitivities in the summer and winter experienced the largest change, with on-road sensitivities becoming less negative in the winter, and less positive in the summer.

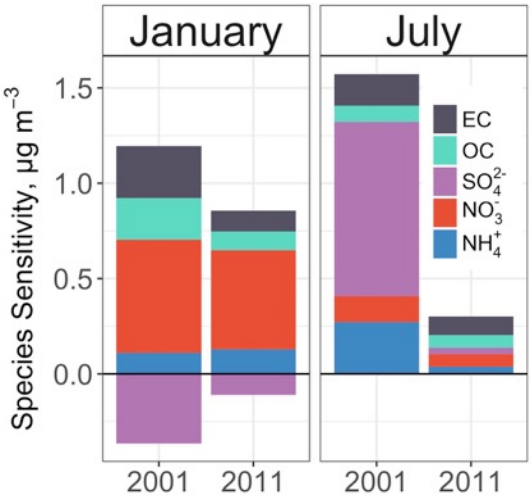


Figure 6-6. Domain-wide average on-road mobile DDM $PM_{2.5}$ species sensitivities.

Average $PM_{2.5}$ sensitivities to EGU emissions tend to be uniformly positive in both winter and summer (Figure 6-7). They are highest near the major sources, e.g., the Ohio River Valley. $PM_{2.5}$ sensitivities to mobile source emissions also tend to be positive, and typically highest in and around major cities, highways and ports (Figure 6-7).

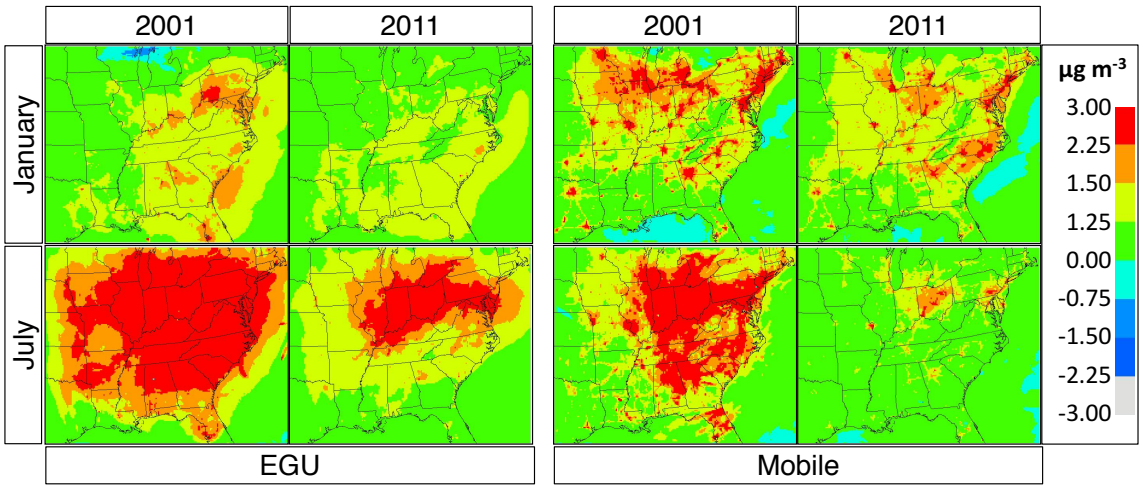


Figure 6-7. Monthly-averaged $PM_{2.5}$ sensitivities to EGU (left) and on-road mobile emissions (right).

Comparisons between sensitivities in the diagnostic CMAQ runs show that pollutant sensitivities have been influenced more by changing emissions than changing meteorology over the decade of interest (Figure 6-8 and Figure 6-9). In January, reduced EGU and mobile emissions are associated with increased (i.e., less negative) sensitivities in rural areas, and little change in within cities. In July, mobile source emission changes are associated with rural reductions in sensitivities and urban increases (because of less O_3 inhibition by NO).

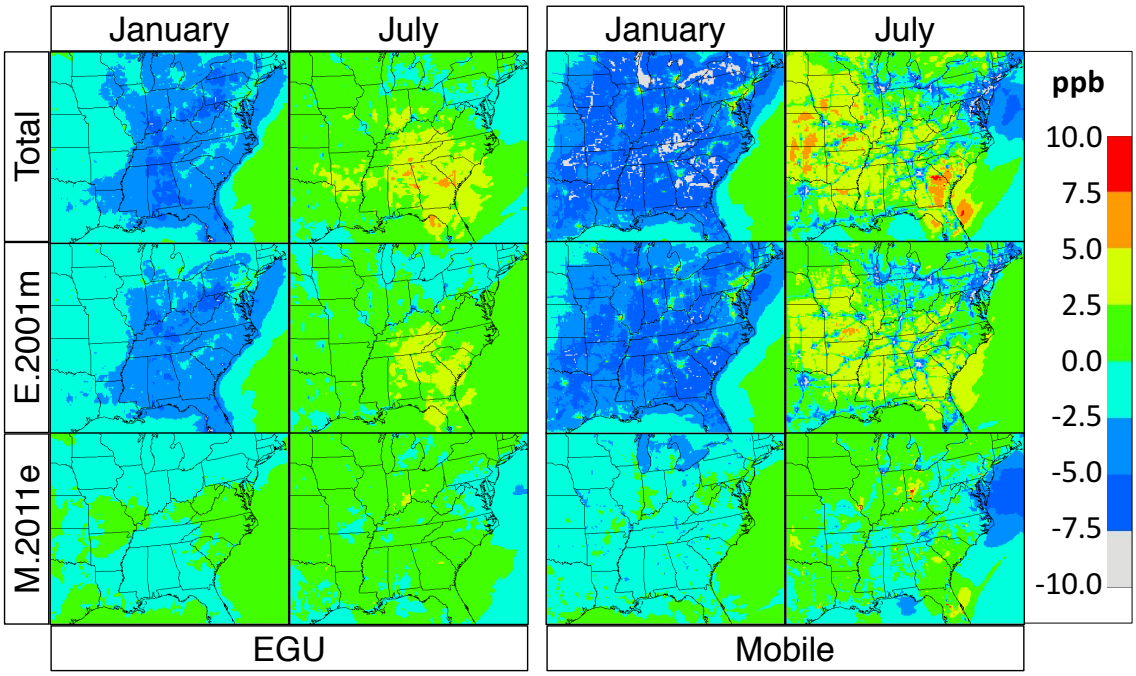


Figure 6-8. Changes in DDM ozone sensitivities to EGU (left) and mobile source (right) emissions attributable to emissions and meteorology changes as described in Table 6-1. Total changes are in the top row, emission-driven changes are in the second row, and meteorology-driven changes are on the bottom row. Positive values denote reduced sensitivity between 2001 and 2011.

January changes in $PM_{2.5}$ sensitivities are much smaller than July changes, and January changes associated with emissions and meteorological changes are comparable in magnitude across most of the country (Figure 6-9). An exception to this is urban sensitivity changes attributable to emissions changes, which have decreased by up to 3 $\mu g\ m^{-3}$ in cities across the domain. In July, both EGU and mobile source emissions reductions led to reduced regional sensitivities.

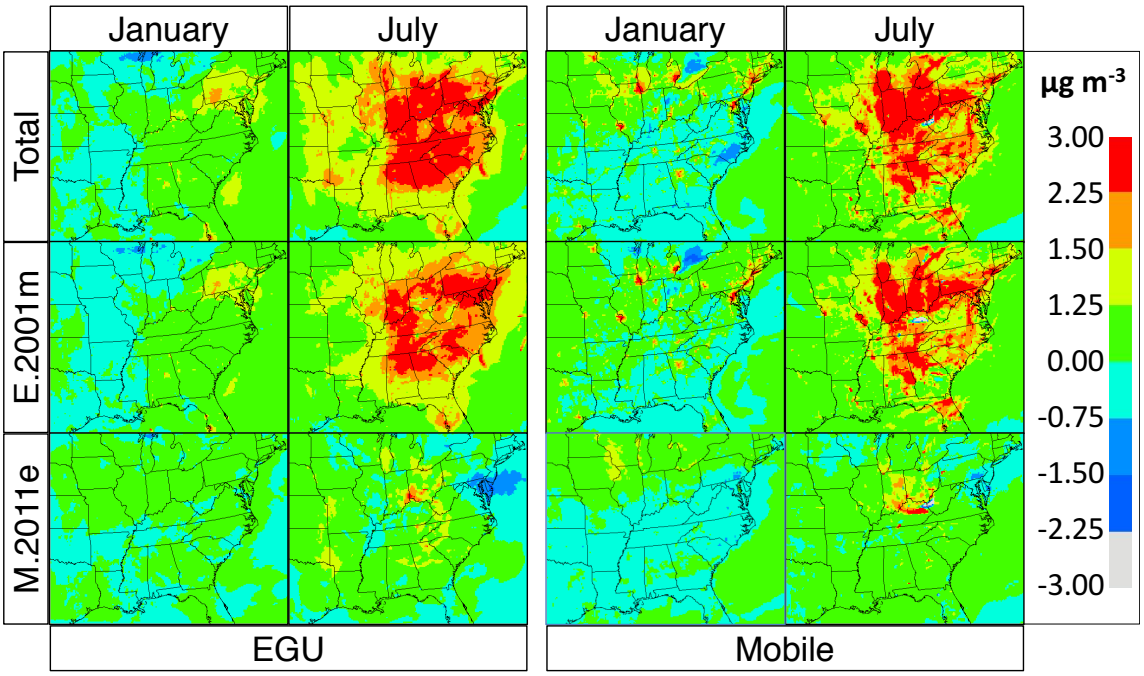


Figure 6-9. Changes in DDM $PM_{2.5}$ sensitivities to EGU (left) and mobile source (right) emissions attributable to emissions and meteorology changes as described in Table 6-1. Total changes are in the top row, emission-driven changes are in the second row, and meteorology-driven changes are on the bottom row. Positive values denote reduced sensitivity between 2001 and 2011.

6.4.3.3 Evaluation of CMAQ and statistical sensitivities

At JST, CMAQ operational performance evaluation yielded results similar to broader evaluations at AQS and SEARCH sites (Table A-1). $PM_{2.5}$ and O_3 concentrations meet at least the criteria benchmark. For O_3, both CMAQ-DDM and empirical sensitivities derived using Eq. 6.11 for EGU and mobile sources are 0.59 (Figure 6-10,Table A-2). $PM_{2.5}$ results overall exhibit less satisfying performance than for ozone, though the CMAQ-DDM and statistical sensitivities comparisons have positive slopes and correlations.

Most of the DDM-statistical sensitivity comparisons exhibit positive slopes and correlations, and species-specific results are reflective of conclusions from the operational evaluation (Table A-2). For example, CMAQ sulfate sensitivities to EGU emissions have a lower magnitude compared to statistical sensitivities (NMB = -16.2%). OC, which was a primary contributor of bias to modeled $PM_{2.5}$ concentrations, yielded low correlations (-0.01 and 0.02) with statistically-modeled EGU and mobile-source sensitivities. EC mobile source sensitivities are highly correlated (r = 0.63) between the two models, though CMAQ sensitivities magnitudes compared to statistical sensitivities. NO_3^- sensitivities exhibit the

poorest fit between the models; CMAQ-DDM EGU sensitivities are lower than the statistical model (-123.4%), and CMAQ-DDM mobile sensitivities are higher (529.6%). For all sensitivity comparisons, the reader is reminded that both approaches have uncertainties, and there is no direct measurement of sensitivities.

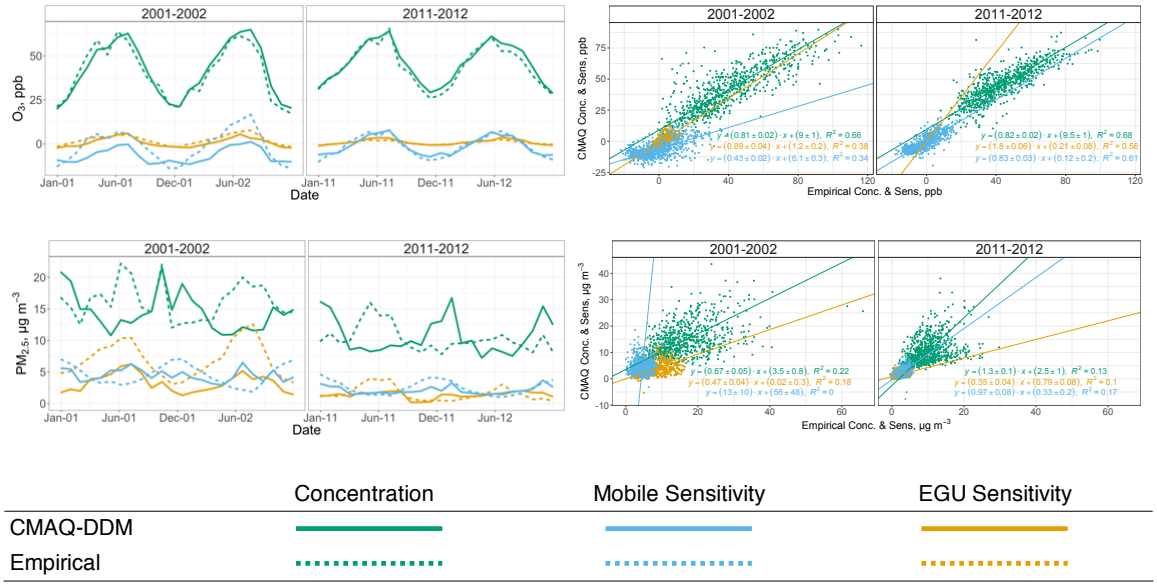


Figure 6-10. Daily O_3 (top) and $PM_{2.5}$ (bottom) CMAQ-statistical comparisons (monthly-averaged time series on the left, daily scatter plot on the right) for ambient concentrations, mobile, and EGU sensitivities. Statistical concentrations (green dashed line) are observations. Sensitivities were compared using orthogonal regression, which assumes error in both variables, whereas concentration comparisons were performed with ordinary least squares regression. Regression fit lines were extended beyond the range of observed data to facilitate the comparisons.

NO_2 and CO modeled concentrations are characterized by large biases and errors and reasonable correlations (0.40 and 0.62, respectively). Compared with statistically-derived sensitivities, CMAQ-derived NO_2 EGU and mobile sensitivities are low and are not strongly correlated. Of the gaseous species besides O_3, mobile CO sensitivities find the greatest agreement between CMAQ and the statistical model. SO_2 concentration and sensitivities are biased low, but the correlations of 0.56, 0.24, and 0.37 are generally higher than the other gaseous species.

6.5 Discussion

This work shows how operational, dynamic, and diagnostic evaluations can expose strengths and weaknesses important in a modeling application tailored to accountability research. The emphasis is broader than traditional model evaluations, which focus on

reproducing observations. Here, more weight is placed on assessing large-scale drivers (emissions and meteorology) and how their variability impacts modeled concentrations. We have assessed these drivers' impacts on modeled concentrations and sensitivities in multiple ways, from isolating their impacts within CMAQ and comparing them to results from multiple empirical statistical models. In the following discussion, we discuss three themes: 1) what the results say about the drivers behind changing air quality in the eastern United States between 2001 and 2012, 2) CMAQ's overall performance in this application and potential sources of bias in the model results, and 3) the application of diagnostic switched-emissions and statistical modeling results as an evaluation tool.

6.5.1 Air Quality in the eastern United States between 2001 and 2012

Observations, along with both PGM and statistical modeling results, show that the United States saw improvements in air quality between 2001 and 2012, and that those improvements are tied to emissions reductions. Observed annual AQS (CMAQ modeled) $PM_{2.5}$ concentrations fell by 4.4 μg m^{-3} (3.8 μg m^{-3}). 2001 saw 34 (23) days on which the domain-wide observed (modeled) MDA8h O_3 average exceeded 60 ppb, and 2012 experienced 8 (5) such days.

For both O_3 and $PM_{2.5}$, multiple deterministic and statistical methods showed that changes over both time (Figure 6-3 and Figure 6-4) and space (Figure 6-5 through Figure 6-9) were largely emissions-driven. In the summertime, the changes in the highest O_3 concentrations between the two 2-year periods were attributable more to meteorology than emissions, though emissions impacts were still large and positive, showing that meteorology plays a major role in peak ozone levels. In the winter, emissions reductions have led almost exclusively to increased concentrations. This effect is observed even in July in city centers, and can be attributed to decreased O_3 production inhibition. Our evaluation and agreement between deterministic and statistical modeling corroborates evidence in Foley et al. (2015b) and elsewhere that, even though the upper end of the O_3 concentration distribution during any one year is highly dependent on meteorological variability, emissions reductions remain an effective mitigation strategy.

Reduced sensitivity magnitudes for O_3 occurred throughout the domain for EGU and mobile sources (Figure 6-5) because emissions were reduced. Sensitivities on a per-unit emissions basis, however, have increased (Figure A-10). Wintertime increases

occurred across the eastern US, though sensitivities in the area surrounding the Ohio River Valley remained large and negative relative to the rest of the domain in 2011. The response in the summertime mirrored that of the winter, with domain-wide sensitivity decreases and continuing importance of major point sources. Ozone sensitivities to on-road mobile source emissions were negative near roadways during the summer, and more negative than the surrounding areas in the wintertime.

Larger reductions in $PM_{2.5}$ emissions sensitivities occurred in the summer than the winter. A reason for this is that changes in wintertime $PM_{2.5}$ between 2001 and 2011 were similarly attributed between emissions and meteorology (Figure 6-4), even with substantial emissions reductions. In the summer, $PM_{2.5}$ reductions were due primarily to emissions reductions. In both EGU and mobile source sensitivities, emissions sensitivities decreases were driven by SO_2 emissions reductions. EGU reductions are associated with the Acid Rain Program, the now-replaced Clean Air Interstate Rule, the Cross-State Interstate Rule, and non-attainment designations under the National Ambient Air Quality Standards. Mobile source emissions reductions of SO_2 were mandated under two national regulatory programs: the Tier II gasoline sulfur program and the 2007 Heavy Duty Diesel Rule (Henneman et al. 2017; Hubbell et al. 2009).

Summertime mobile sensitivities of SO_4^{2-} and NH_4^+ levels—both of which are enhanced by SO_2 in the atmosphere—decreased by 96% and 86%, respectively. Mobile source sensitivities of NO_3^-, which is an important winter time $PM_{2.5}$ constituent, are largely unchanged from 2001-2011. Further, mobile emissions changes have led to decreased negative SO_4^{2-} sensitivity, meaning emissions reductions have led to slightly increased SO_4^{2-} wintertime concentrations due to less competition for hydroxyl radical. Wintertime $PM_{2.5}$ benefits of emissions reductions were primarily achieved in the immediate (>100km) vicinity of EGU sources. In the summer, changes are more regional, when transport and secondary chemistry is more important.

6.5.2 Model performance evaluation

Here, we summarize the major findings for questions posed for the operational, dynamic, and diagnostic evaluations at the end of the introduction regarding the evaluation of the modeling setup. O_3, $PM_{2.5}$, and each $PM_{2.5}$ species all meet one or more criteria benchmark defined by Emery et al. (2017) per modeled year. O_3 and $PM_{2.5}$—the pollutants

of primary regulatory interest and for which PGM models are most often applied—met at least the criteria benchmark for all metrics in all four years. The dynamic evaluation found that CMAQ generally captured the changing concentrations of O_3 and $PM_{2.5}$.

Since the metric benchmarks were designed for regional spatial scales or below and time scales of seasonal or less, our application to annual runs over the eastern United States is slightly outside the original recommended scales. To investigate impacts of applying the metrics, we performed an operational analysis for O_3 and $PM_{2.5}$ on two sub-domains, above ("North") and below ("South") 37º latitude (Table A-4). Differences do exist between the two domains; for example, $PM_{2.5}$ is negatively biased in all years in the South and positively biased in the North. The overall picture related to the benchmarks (and, more importantly, the broader assertion that the model results capture reality at a level consistent with previous applications), however, is largely unchanged.

In the dynamic analysis, NMB proved a useful tool to assess CMAQ's ability to capture changes over the decade. NME, however, was high for all species, even those that met criteria metrics for other evaluation statistics, due to the large daily variability in the sign of the change in simulated concentrations versus the observed changes.

6.5.2.1 Ozone

Evaluated without a cutoff, simulated hourly and MDA8h O_3 are positively biased, and are negatively biased with a 40 ppb cutoff. These results agree with other modeling studies in the Eastern United States (Fiore et al. 2009; Foley et al. 2015b; Reidmiller et al. 2009). Travis et al. (2016), using evidence from other studies that mobile and industrial source NO_X in the NEI is biased high by up to 60%, found a better agreement with observations using reduced NO_X emissions. We investigated model performance using one additional simulation for July 2011 using mobile NO_X decreased by 50%. Results show that simulated July hourly ozone is negatively biased throughout the domain, and reducing mobile NO_X leads to an increased negative bias (Table A-5). MDA8h O_3, on the other hand, is biased high in the base case, and reducing mobile NO_X emissions reduces this bias. Overall, the results show that reducing mobile NO_X emissions decreases CMAQ simulated O_3, though given the likelihood of other issues in the modeling framework, this comparison is likely not indicative of the accuracy of either inventory.

CMAQ-modeled O_3 sensitivities are lower than those derived statistically, but highly correlated with other species compared to statistically-modeled sensitivities at JST in Atlanta (Table A-2). The lower simulated sensitivities are suggestive of a potential high bias in NO_X emissions estimates, since, in the city center, increased emissions have a negative impact on O_3 concentrations except when O_3 levels are high.

Friberg et al. (2016) showed that CMAQ is better able to capture temporal variability than complex spatial patterns. Results of this modeling effort yield similar conclusions; average temporal correlations for both hourly and daily O_3 are higher than spatial correlations (Table A-6).

6.5.2.2 $PM_{2.5}$ and $PM_{2.5}$ species

The results of the operational and dynamic evaluations suggest that changing bias throughout the year in $PM_{2.5}$ species in different directions aids in the simulation of total $PM_{2.5}$. For instance, negative OC and SO_4^{2-} biases in the summertime are balanced by positive EC bias. Similarly, high bias in modeled OC and SO_4^{2-} change across the decade compensate for low bias in modeled EC change.

OC exhibits low bias and error, but low correlations and variable MB across the year (Figure 6-1) show that it failed to capture variability throughout the year. In the summer, this result reflects the well-documented incompleteness of CMAQ in capturing secondary organic aerosol chemistry (Napelenok et al. 2014; Pye et al. 2015). In the winter, mischaracterization of biomass burning emissions contributed bias to OC modeled concentrations (Hu et al. 2014). Continued improvements in the characterization of chemistry, particularly SOA formation, should also help reduce difference between simulations and observations (Glasius and Goldstein 2016; Goldberg et al. 2016). The upcoming version of CMAQ is expected to address these biases by improving interactions between atmospheric carbon and aerosol liquid water (Pye et al. 2016).

In the summer, both CMAQ-modeled SO_4^{2-} concentrations and changes over the decade are biased high. CMAQ attributes much of the change to mobile SO_4^{2-} sensitivities changes over the decade, which is not reflected in the statistical modeling. These results may indicate the inability of the statistical modeling to capture secondary SO_4^{2-} chemistry from mobile sources; however, it is difficult to draw strict conclusions in a comparison of two models.

The EC operational evaluation yielded correlations that meet the criteria benchmark and relatively high NMB and NME. The change in EC is biased low across the decade. This suggests that NEI emissions of EC, a primary pollutant whose source is automobiles, capture daily variability, but overestimate the magnitude and change across years.

Spatial and temporal $PM_{2.5}$ are much more similar than for O_3 (Table A-6). For 3 of 4 modeled years, temporal correlations exceed spatial correlations, but the differences are small.

6.5.3 Diagnostic model evaluation tools

The diagnostic evaluation portion of this work makes novel comparisons between meteorology and emissions-influenced concentration changes modeled with CMAQ-DDM and ones inferred from statistical modeling of observations. An evaluation in this vein is particularly relevant in PGM applications for accountability research, since meteorological fluctuations are a major contributor to uncertainty in attributing causality of air quality and health outcomes changes to regulatory actions.

Differences in meteorological and emissions-influenced changes across season and pollutant have important consequences on future policy—small $PM_{2.5}$ changes attributable to emissions or meteorology (reinforced by small sensitivity changes) suggests that a future approach for reducing wintertime $PM_{2.5}$ concentrations is reducing NO_X emissions. This finding adds to the complexity of regulatory decisions, since nitrate concentrations have been shown to be less sensitive to emissions reductions reducing NO_X emissions (Blanchard and Hidy 2003); further, wintertime NO_X emission reductions would have the effect of increasing ozone, particularly in urban areas.

Median meteorological impacts on both O_3 and $PM_{2.5}$ were near zero in both January and July across the SEARCH sites, and median emissions impacts were positive for O_3 in July and $PM_{2.5}$ in both months (Figure 6-3 and Figure 6-4), suggesting that emissions controls were effective at reducing concentrations on most days and locations throughout the Southeast in July. The agreement between CMAQ-modeled and statistical meteorological and emissions contributions provides evidence that the estimated impacts reflect reality.

A second variety of diagnostic test assessed CMAQ-modeled sensitivities to emissions using statistically-modeled sensitivities. This analysis showed that the two

models produce similar sensitivities for O_3 and $PM_{2.5}$, but the models agree less for other gases and $PM_{2.5}$ species. We compared sensitivities by scaling them by the ratio of observed to CMAQ-simulated concentrations. This generally improved fit in ozone and $PM_{2.5}$, but led to a negative correlation between mobile $PM_{2.5}$ sensitivities. This occurred because CMAQ-simulated mobile sensitivities are smaller in the wintertime than statistical, and vice versa in the summer. Further, CMAQ concentrations are biased high in the summer and low in the winter. Correcting sensitivities for concentration bias creates anti-correlated sensitivities.

Each of the models used here (CMAQ, detrending, and statistical emissions-concentrations models) comes with its own set of assumptions and uncertainties. The consistency between results of the methods, however, provides evidence that they capture the atmospheric dynamics important to the current application to accountability assessment, though performance differs based on the pollutant.

6.6 Conclusions

Detailed model analysis using traditional and novel methods furthers our understanding of effects of air quality regulations across the eastern United States. In the operational evaluation, CMAQ simulated most the species at accuracy levels consistent with previous modeling approaches. In the dynamic evaluation, CMAQ more accurately simulated O_3 and $PM_{2.5}$ changes over time compared to $PM_{2.5}$ species, suggesting that biases in the modeled species concentrations canceled each other out. In the diagnostic evaluation, CMAQ and statistically-derived sensitivities to emissions and meteorology agree in the southeast, and suggest that emissions reductions have led to a shrinking O_3 distribution and have reduced $PM_{2.5}$ concentrations across its distribution.

Major reductions of multiple air pollutants have largely been due to emissions reductions (and previous work has attributed these emissions reductions to regulations). Epidemiologic and health impact analyses, presented in a report to the Health Effects Institute (Russell et al., 2017) and a companion manuscript (Abrams et al., n.d.) further links those air quality changes to health benefits. Further, the analysis shows that EGU and mobile source emissions reductions remain a viable option for improving air quality; however, in the wintertime, emissions reductions of NO_X (SO_2) were linked with increasing O_3 ($PM_{2.5}$).

Our method employed two types of models: deterministic and statistical, both of which provide important information. In areas that the two models agreed (for instance, in the attribution of variability to emissions and meteorology), the results increase confidence that the findings are accurate. Other areas, such as model sensitivities, require more model development and analysis, reflecting that sensitivities are not directly observable.

6.7 Acknowledgements

This material is based upon work supported by Health Effects Institute and the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1148903. We thank Naresh Kumar of the Electric Power Research Institute (EPRI) for helpful comments. The SEARCH network was funded by Southern Company and EPRI, and we thank Eric Edgerton for his help and useful discussions.

6.8 References

Abrams J, Klein M, Henneman LR, Russell AG, Sarnat SE, Chang HH, et al. Impacts of air pollution control policies on cardiorespiratory emergency department visits, Atlanta, GA 1999-2013.

Banzhaf S, Schaap M, Kranenburg R, Manders AMM, Segers AJ, Visschedijk AJH, et al. 2015. Dynamic model evaluation for secondary inorganic aerosol and its precursors over Europe between 1990 and 2009. Geosci. Model Dev. 8:1047–1070; doi:10.5194/gmd-8-1047-2015.

Blanchard CL, Hidy GM, Tanenbaum S. 2014. Ozone in the southeastern United States: An observation-based model using measurements from the SEARCH network. Atmos. Environ. 88:192–200; doi:10.1016/j.atmosenv.2014.02.006.

Blanchard CL, Hidy GM, Tanenbaum S, Edgerton ES, Hartsell BE. 2013. The Southeastern Aerosol Research and Characterization (SEARCH) study: Spatial variations and chemical climatology, 1999–2010. J. Air Waste Manage. Assoc. 63:260–275; doi:10.1080/10962247.2012.749816.

Boylan JW, Russell AG. 2006. PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models. Atmos. Environ. 40:4946–4959; doi:10.1016/j.atmosenv.2005.09.087.

Bravo MA, Fuentes M, Zhang Y, Burr MJ, Bell ML. 2012. Comparison of exposure estimation methods for air pollutants: Ambient monitoring data and regional air quality simulation. Environ. Res. 116:1–10; doi:10.1016/j.envres.2012.04.008.

Brune WH, Baier BC, Thomas J, Ren X, Cohen RC, Pusede SE, et al. 2016. Ozone production chemistry in the presence of urban plumes. Faraday Discuss.; doi:10.1039/C5FD00204D.

Byun D, Schere KL. 2006. Review of the Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multiscale Air Quality (CMAQ) Modeling System. Appl. Mech. Rev. 59:51; doi:10.1115/1.2128636.

Camalier L, Cox W, Dolwick P. 2007. The effects of meteorology on ozone in urban areas and their use in assessing ozone trends. Atmos. Environ. 41:7127–7137; doi:10.1016/j.atmosenv.2007.04.061.

Cohan DS. 2004. Applicability of CMAQ-DDM to Source Apportionment and Control Strategy Development. 3rd Annu. C. Model. Users' Conf.

Cox WM, Chu S-H. 1993. Meteorologically adjusted ozone trends in urban areas: a probabilistic approach. Atmos. Environ. 27:425–434; doi:doi:10.1016/0957-1272(93)90019-3.

Daskalakis N, Tsigaridis K, Myriokefalitakis S, Fanourgakis GS, Kanakidou M. 2016. Large gain in air quality compared to an alternative anthropogenic emissions scenario. Atmos. Chem. Phys. 16:9771–9784; doi:10.5194/acp-16-9771-2016.

Dennis R, Fox T, Fuentes M, Gilliland A, Hanna S, Hogrefe C, et al. 2010. A Framework for Evaluating Regional-Scale Numerical Photochemical Modeling Systems. Environ. fluid Mech. 10:471–489; doi:10.1007/s10652-009-9163-2.

Di Q, Kloog I, Koutrakis P, Lyapustin A, Wang Y, Schwartz J. 2016. Assessing PM2.5 Exposures with High Spatiotemporal Resolution across the Continental United States. Environ. Sci. Technol. 50:4712–4721; doi:10.1021/acs.est.5b06121.

Emery C, Liu Z, Russell A, Talat Odman M, Yarwood G, Kumar N. 2016. Recommendations on Statistics and Benchmarks to Assess Photochemical Model Performance. J. Air Waste Manage. Assoc.

Fann N, Lamson AD, Anenberg SC, Wesson K, Risley D, Hubbell BJ. 2012. Estimating the national public health burden associated with exposure to ambient PM2.5 and ozone. Risk Anal. 32:81–95; doi:10.1111/j.1539-6924.2011.01630.x.

Fiore AM, Dentener FJ, Wild O, Cuvelier C, Schultz MG, Hess P, et al. 2009. Multimodel estimates of intercontinental source-receptor relationships for ozone pollution. J. Geophys. Res. Atmos. 114:1–21; doi:10.1029/2008JD010816.

Foley KM, Dolwick P, Hogrefe C, Simon H, Timin B, Possiel N. 2015a. Dynamic evaluation of CMAQ part II: Evaluation of relative response factor metrics for ozone attainment demonstrations. Atmos. Environ. 103:188–195; doi:10.1016/j.atmosenv.2014.12.039.

Foley KM, Hogrefe C, Pouliot G, Possiel N, Roselle SJ, Simon H, et al. 2015b. Dynamic evaluation of CMAQ part I: Separating the effects of changing emissions and changing meteorology on ozone levels between 2002 and 2005 in the eastern US. Atmos. Environ. 103:247–255; doi:10.1016/j.atmosenv.2014.12.038.

Foley KM, Roselle SJ, Appel KW, Bhave P V., Pleim JE, Otte TL, et al. 2010. Incremental testing of the community multiscale air quality (CMAQ) modeling system version 4.7. Geosci. Model Dev. 3:205–226; doi:10.5194/gmd-3-205-2010.

Gégo E, Gilliland A, Godowitch J, Rao S, Porter P, Hogrefe C. 2008. Modeling Analyses of the Effects of Changes in Nitrogen Oxides Emissions from the Electric Power Sector on Ozone Levels in the Eastern United States. J. Air Waste Manage. Assoc. 58:580–588; doi:10.3155/1047-3289.58.4.580.

Gilliland AB, Hogrefe C, Pinder RW, Godowitch JM, Foley KL, Rao ST. 2008. Dynamic evaluation of regional air quality models: Assessing changes in O3 stemming from changes in emissions and meteorology. Atmos. Environ. 42:5110–5123; doi:10.1016/j.atmosenv.2008.02.018.

Glasius M, Goldstein AH. 2016. Recent Discoveries and Future Challenges in Atmospheric Organic Chemistry. Environ. Sci. Technol. 50:2754–2764; doi:10.1021/acs.est.5b05105.

Godowitch JM, Gilliland a. B, Draxler RR, Rao ST. 2008. Modeling assessment of point source NOx emission reductions on ozone air quality in the eastern United States. Atmos. Environ. 42:87–100; doi:10.1016/j.atmosenv.2007.09.032.

Godowitch JM, Pouliot G a., Trivikrama Rao S. 2010. Assessing multi-year changes in modeled and observed urban NOX concentrations from a dynamic model evaluation perspective. Atmos. Environ. 44:2894–2901; doi:10.1016/j.atmosenv.2010.04.040.

Goldberg DL, Vinciguerra TP, Anderson DC, Hembeck L, Canty TP, Ehrman SH, et al. 2016. CAMx Ozone Source Attribution in the Eastern United States using Guidance from Observations during DISCOVER-AQ Maryland. Geophys. Res. Lett. 2249–2258; doi:10.1002/2015GL067332.

Hakami A, Odman T, Russell AG. 2004. Nonlinearity in atmospheric response: A direct sensitivity analysis approach. J. Geophys. Res. 109:D15303; doi:10.1029/2003JD004502.

Health Effects Institute. 2003. Assessing Health Impact of Air Quality Regulations: Concepts and Methods for Accountability Research.

Henneman LR, Chang HH, Lavoue D, Mulholland JA, Russell AG. 2017. Accountability assessment of regulatory impacts on ozone and PM2.5 concentrations using statistical and deterministic pollutant sensitivities. Air Qual. Atmos. Heal.; doi:10.1007/s11869-017-0463-2.

Henneman LR, Holmes HA, Mulholland JA, Russell AG. 2015. Meteorological Detrending of Primary and Secondary Pollutant Concentrations: Method Application and Evaluation Using Long-Term (2000-2012) Data in Atlanta. Atmos. Environ. 119:201–210; doi:10.1016/j.atmosenv.2015.08.007.

Hogrefe C, Pouliot G, Wong D, Torian A, Roselle S, Pleim J, et al. 2015. Annual application and evaluation of the online coupled WRF-CMAQ system over North America under AQMEII phase 2. Atmos. Environ. 115:683–694; doi:10.1016/j.atmosenv.2014.12.034.

Hu J, Chen J, Ying Q, Zhang H. 2016. One-year simulation of ozone and particulate matter in China using WRF/CMAQ modeling system. Atmos. Chem. Phys. 16:10333–10350; doi:10.5194/acp-16-10333-2016.

Hu Y, Balachandran S, Pachon JE, Baek J, Ivey C, Holmes H, et al. 2014. Fine particulate matter source apportionment using a hybrid chemical transport and receptor model approach. Atmos. Chem. Phys. 14:5415–5431; doi:10.5194/acp-14-5415-2014.

Hubbell BJ, Crume R V., Evarts DM, Cohen JM. 2009. Policy Monitor: Regulation and Progress under the 1990 Clean Air Act Amendments. Rev. Environ. Econ. Policy 4:122–138; doi:10.1093/reep/rep019.

Kuebler J, van den Bergh H, Russell AG. 2001. Long-term trends of primary and secondary pollutant concentrations in Switzerland and their response to emission controls and economic changes. Atmos. Environ. 35:1351–1363; doi:10.1016/S1352-2310(00)00401-5.

Liu XH, Zhang Y, Xing J, Zhang Q, Wang K, Streets DG, et al. 2010. Understanding of regional air pollution over China using CMAQ, part II. Process analysis and sensitivity of ozone and particulate matter to precursor emissions. Atmos. Environ. 44:3719–3727; doi:10.1016/j.atmosenv.2010.03.036.

Muller N, Tong D, Mendelsohn RO. 2009. Regulating NOx and SO2 Emissions in Atlanta. BE J. Econ. Anal. Policy Contrib. to Econ. Anal. Policy 9.

Napelenok SL, Foley KM, Kang D, Mathur R, Pierce T, Rao ST. 2011. Dynamic evaluation of regional air quality model's response to emission reductions in the presence of uncertain emission inventories. Atmos. Environ. 45:4091–4098; doi:10.1016/j.atmosenv.2011.03.030.

Napelenok SL, Simon H, Bhave P V., Pye HOT, Pouliot GA, Sheesley RJ, et al. 2014. Diagnostic air quality model evaluation of source-specific primary and secondary fine particulate carbon. Environ. Sci. Technol. 48:464–473; doi:10.1021/es403304w.

National Research Council. 2004. *Air Quality Management in the United States*. National Academies Press.

Odman MT, Hu Y, Chang ME, Russell AG. 2007. Chapter 2.14 Forecasting ozone and PM2.5 in southeastern U.S. Dev. Environ. Sci. 6:220–229; doi:10.1016/S1474-8177(07)06214-6.

Park RJ, Jacob DJ, Field BD, Yantosca RM, Chin M. 2004. Natural and transboundary pollution influences on sulfate-nitrate-ammonium aerosols in the United States: Implications for policy. J. Geophys. Res. D Atmos. 109; doi:10.1029/2003JD00473.

Park S-K, Evan Cobb C, Wade K, Mulholland J, Hu Y, Russell AG. 2006. Uncertainty in air quality model evaluation for particulate matter due to spatial variations in pollutant concentrations. Atmos. Environ. 40:563–573; doi:10.1016/j.atmosenv.2005.11.078.

Pye HOT, Luecken DJ, Xu L, Boyd CM, Ng NL, Baker KR, et al. 2015. Modeling the Current and Future Roles of Particulate Organic Nitrates in the Southeastern United States. Environ. Sci. Technol. 49:14195–14203; doi:10.1021/acs.est.5b03738.

Pye HOT, Murphy BN, Xu L, Ng NL, Carlton AG, Guo H, et al. 2016. On the implications of aerosol liquid water and phase separation for organic aerosol mass. Atmos. Chem. Phys. Discuss. 1–44; doi:10.5194/acp-2016-719.

Rao ST, Galmarini S, Puckett K. 2011. Air quality model evaluation international initiative (AQMEII): Advancing the state of the science in regional photochemical modeling and its applications. Bull. Am. Meteorol. Soc. 92:23–30; doi:10.1175/2010BAMS3069.1.

Rao STS, Zurbenko IG, Neagu R, Porter PS, Ku JY, Henry RF. 1997. Space and Time Scales in Ambient Ozone Data. Bull. Am. Meteorol. Soc. 78:2153–2166; doi:10.1175/1520-0477(1997)078<2153:SATSIA>2.0.CO;2.

Reidmiller DR, Fiore AM, Jaffe DA, Bergmann D, Cuvelier C, Dentener FJ, et al. 2009. The influence of foreign vs. North American emissions on surface ozone in the US. Atmos. Chem. Phys. 9:5027–5042; doi:10.5194/acpd-9-7927-2009.

Russell, Armistead G, Paige E Tolbert, Lucas RF Henneman, Joe Abrams, Cong Liu, Mitch Klein, James A Mulholland, et al. 2017. Impacts of Regulations on Air Quality and Emergency Department visits in the Atlanta metropolitan area, 1999-2013. *In press*.

Simon H, Baker KR, Akhtar F, Napelenok SL, Possiel N, Wells B, et al. 2013. A direct sensitivity approach to predict hourly ozone resulting from compliance with the national ambient air quality standard. Environ. Sci. Technol. 47:2304–2313; doi:10.1021/es303674e.

Simon H, Baker KR, Phillips S. 2012. Compilation and interpretation of photochemical model performance statistics published between 2006 and 2012. Atmos. Environ. 61:124–139; doi:10.1016/j.atmosenv.2012.07.012.

Simon H, Reff A, Wells B, Xing J, Frank N. 2014. Ozone trends across the United States over a period of decreasing NOx and VOC emissions. Environ. Sci. Technol. 49:186–95; doi:10.1021/es504514z.

Souri AH, Choi Y, Jeon W, Li X, Pan S, Diao L, et al. 2016. Constraining NOx emissions using satellite NO2 measurements during 2013 DISCOVER-AQ Texas campaign. Atmos. Environ. 131:371–381; doi:10.1016/j.atmosenv.2016.02.020.

Travis KR, Jacob DJ, Fisher JA, Kim PS, Marais EA, Zhu L, et al. 2016. Why do models overestimate surface ozone in the Southeast United States? Atmos. Chem. Phys. 16:13561–13577; doi:10.5194/acp-16-13561-2016.

U.S. EPA. 2014. Draft Modeling Guidance for Demonstrating Attainment of Air Quality Goals for Ozone PM2.5, and Regional Haze.

U.S. EPA. 1991. Guideline for regulatory application of the urban airshed model.; doi:EPA-450/4-91-013.

U.S. EPA. 2015a. National Ambient Air Quality Standards for Ozone Final Rule. 40 CFR Parts 50,51,52,53 58 80: 1–7.

U.S. EPA. 2005. Regulatory Impact Analysis for the Final Clean Air Interstate Rule.

U.S. EPA. 1998. Regulatory Impact Analysis for the NBP.

U.S. EPA. 2015b. Technical Support Document EPA's 2011 National-scale Air Toxics Assessment 2011.

Xing J, Mathur R, Pleim J, Hogrefe C, Gan CM, Wong DC, et al. 2015. Can a coupled meteorology-chemistry model reproduce the historical trend in aerosol direct radiative effects over the Northern Hemisphere? Atmos. Chem. Phys. 15:9997–10018; doi:10.5194/acp-15-9997-2015.

CHAPTER 7. OZONE RESPONSES TO RECENT AND FUTURE EMISSIONS CHANGES IN THE EASTERN UNITED STATES

7.1 Abstract

Ozone production efficiency (OPE), a measure of the number of ozone (O_3) molecules produced per emitted NO_X ($NO + NO_2$) molecule, helps establish the relationship between NOx emissions and ozone formation. We estimate OPE in the eastern United States using two novel approaches: a spline-based empirical method using observations at eight measurement sites and a chemical transport model (CTM) method based on concentration and depositions sensitivities. The CTM approach allows for the explicit control of differing deposition rates in O_3 and NO_X reaction products (NO_Z), and can estimate on-road mobile and electricity generating unit source OPEs. We find that OPE has increased throughout the United States, but is bounded by location-specific upper limits at low NO_Z. The CTM and observation-based approaches agree at low NO_Z concentrations, but CTM OPEs are greater than statistically-calculated OPEs at high NO_Z, suggesting a greater importance of deposition differences between O_3 and NO_Z at high NO_Z concentrations. Results suggest that NO_X emissions reductions become more effective at lower NO_Z concentrations, but below location-specific NO_Z concentrations O_3 responses become linear. Electricity generating unit OPEs are higher than mobile OPEs except near stacks, meaning additional utility emissions reductions will have greater per-unit impacts than mobile source emissions reductions.

7.2 Introduction

Atmospheric scientists have understood the underlying mechanisms of ozone (O_3) production for decades; however, air quality managers continue to have trouble meeting ozone standards, particularly as they have been tightened in response to health studies identifying impacts at lower concentrations (Simon et al. 2014; U.S. EPA 2015). The availability of detailed ambient monitoring data and much expanded modeling capabilities presents an opportunity for a detailed assessment of the effectiveness of past and future air quality improvement efforts. An assessment of this type—investigating changing mechanisms attributable to air quality controls—serves to link accountability assessments

(i.e., efforts to quantify the impacts of existing regulatory programs) to proposed air quality programs.

Largely as a result of air quality management policies, anthropogenic emissions of two contributors to ozone concentrations—nitrogen oxides ($NO_X = NO + NO_2$) and volatile organic compounds (VOCs)—have decreased throughout the eastern United States over the past two decades, particularly in urban areas (Vijayaraghavan et al. 2014; Xing et al. 2013). Emissions reductions, particularly in electrical generating unit (EGU) and mobile source sectors, have led to reduced levels of ambient NO_X, NO_X reaction products (NO_Z = total atmospheric oxidized reactive nitrogen (NO_Y) minus NO_X), and VOCs (Hidy and Blanchard 2015; He et al. 2013). These reductions have been linked to changing O_3 concentrations, including reductions (primarily in the summer) and increases (primarily in the winter and VOC limited urban areas) (Simon et al. 2014). Multiple studies have found that meteorology has a major impact on daily O_3 concentrations (Henneman et al. 2015; Camalier, Cox, and Dolwick 2007; Hidy and Blanchard 2015), and climate variability and synoptic patterns impact longer-term trends (L. Shen, Mickley, and Tai 2015; Lu Shen and Mickley 2017).

Ozone production efficiency (OPE) has historically been a popular method for reporting the relationship between O_3 and NO_X (Seinfeld and Pandis 2006). Seinfeld and Pandis (2006) define OPE as "the number of molecules of O_3 formed per each NO_X molecule removed from the system" (Seinfeld and Pandis 2006). NO_X generates O_3 as it cycles between NO and NO_2, and since each NO_X molecule that enters the system must leave, the emissions rate approximates the loss rate. The reaction of NO_2 with OH—which forms nitric acid (HNO_3)—is typically the dominant pathway for NO_X loss, though it can combine with organics and form HNO_3 via N_2O_5 hydrolysis (which occurs primarily at night). OPE has real policy implications, since it provides an estimate for how much less ozone would be formed for a specified reduction in NO_X emissions (Liu et al. 1987; Ryerson et al. 1998).

Previous studies using both observation-based and chemical transport model (CTM)-based approaches have found increasing OPE as NO_X or NO_Z decreases (Kleinman et al. 2002; Liu et al. 1987). This increase carries the implication that NO_X emissions reductions become increasingly effective. There has, however, been little investigation of

OPE at very low NO_Z concentrations. We set out to fulfil two objectives: 1) assess multiple observation and CTM-based approaches for estimating total and source-specific OPE in the eastern United States, and 2) project impacts of continuing NO_X emissions reductions on future OPEs and O_3 concentrations.

7.3 Methods

7.3.1 Observation-based OPE

We use observations from the SouthEastern Aerosol Characterization (SEARCH) network sampling sites (Hansen et al. 2003; Hansen et al. 2006), which provide meteorological data, O_3, NO_X, and NO_Y species observations at eight locations from 1996-2015 (not all sites were online for the entire period). The locations consist of four pairs: three urban-rural and one urban-suburban. Throughout this manuscript, we refer to the sites by their three-letter acronym, including a superscript that denotes if each site is in an Urban (U), Rural (R) or Suburban (S) location. The southeastern United States is characterized by hot, humid summers with periods of stagnation interrupted by large thunderstorms. Winters are typically mild and drier, though extended periods of wet weather accompany frontal activity. Isolated city centers and power plants, situated throughout mixed coniferous and deciduous forests and agricultural land, serve as major air pollution sources (Hansen et al. 2003). Biomass burning, primarily in the winter, contributes to air pollution.

Trainer et al. (1993) proposed the relationship between O_3 and NO_Z as an empirical estimate of OPE; in a single plume, the loss of NO_X equals the production of NO_Z when the deposition rate of NO_X is negligible (Thornton et al. 2002). Kleinman et al. (2002) noted that a major limitation to using a fixed-site monitor to estimate OPE is that air masses crossing a monitor have potentially very different histories (Kleinman et al. 2002). Take, for instance, the case of a monitor near a city: at times, it would be downwind of the city; O_3 (and NO_Z) levels would be high as the city contributes fresh emissions that lead to high ozone levels, but the NO_Z would have had little time to deposit. At other times, the monitor is upwind, the air masses impacting the monitor have lower ozone, and much of the NO_Z will have been lost to deposition. Relating measured O_3 to NO_Z in this case is not indicative of the OPE.

We controlled for air parcel history by stratifying days based on their photochemical activity. Days with high photochemical activity have similar meteorological

characteristics, meaning changes in ozone levels should relate to changes in NO_Z, though other emissions changes can also play a role. We used long term meteorological detrending of ozone observations to identify the 20% of days with all the related data available at each site that were most photochemically active. Photochemical state (PS*, described in detail in Henneman et al., 2017) was estimated using the meteorological detrending method presented in Henneman et al. (2015). In brief, observed O_3 was split into five trends with different periods: long-term (period > 365 days), seasonal (365 days), week-holiday (7-365 days), short-term meteorological (< 365 days), and white-noise (1 day). The splitting was accomplished using a combination of statistical techniques, including KZ filters (Zurbenko 1991) and regressions with meteorological variables. PS* is the combination of the seasonal and short-term meteorological trends, and represents an emissions-independent estimate of daily atmospheric activity.

We used 2-3 pm averages (local time) for daily metrics of all observed species throughout this work, and assessed only days in the upper 20th percentile for PS*. NO_Z was calculated as the difference between NO_Y and NO_X. We tested multiple models for this relationship, including linear and log models. Models of these types, however, are limited in their application to very low NO_Z values, where the slope ($d[O_3]/d[NO_Z]$) increases and data is sparse. Spline models divide the domain of the independent variable into sections, and fit curves to each section (Hastie et al. 2009). The sections are divided at 'knots', and models fit in each of the sections are required to be continuous at the knots. We applied cubic splines, which extend the constraints to equal first and second derivatives at each knot, using the *rcs* and *ols* commands from the *rms* package in R (version 3.3.3) (R Development Core Team 2012). We take the marginal OPE—i.e., the additional amount of ozone formed due to an additional amount of NO_Z—as the derivative of fitted O_3 vs. NO_Z models at each site. The spline model captures slope changes at multiple points and does not result in nonphysically high slopes at low NO_Z (which are found using the log model). Slopes (i.e., OPEs) are constant outside the outermost knots.

7.3.2 Chemical Transport Model-based OPE

We used the Community Multiscale Air Quality (CMAQ) model with coupled direct decoupled method (DDM) (Byun and Schere 2006) to simulate air quality and first-order sensitivities to emissions over a 12 km domain covering the eastern United States.

The model setup and application was first described in Henneman et al. (2017) and evaluated in detail in Henneman et al., n.d. We used results from the 2001 and 2011 simulations described in our previous work along with two additional month-long cases with July 2011 meteorology and domain-wide NO_X emissions reduced by 50% and 90%. The latter simulations serve to control for meteorology and all other emissions sources completely while reducing NO_X emissions. We preserve consistency with observations by using 2-3 p.m. averages, and focus on July because it is a hot summer month with typical conditions suitable for O_3 formation.

For a first estimate of CTM-simulated OPE, we used a brute force (BF) method that assumes linear OPEs across changing NO_Z concentrations similar to previous approaches (Trainer et al. 1993; Godowitch, Gilliam, and Rao 2011):

$$OPE^{BF} = \frac{\Delta O_3}{\Delta NO_Z} \quad \text{(Eq. 7.1)}$$

Where NO_Z is the sum of NO_X reaction products—peroxyacetyl nitrate (PAN and PANX), peroxynitric acid (PNA), organic nitrate (NTR), nitric acid (HNO3), nitrous acid (HONO), nitrate radical (NO3), and dinitrogen pentoxide (N2O5)—in the CBO5 chemical mechanism as N (e.g., N_2O_5 contains two N molecules). The changes in O_3 and NO_Z are matched by site and date across CMAQ runs.

We compare OPE^{BF} with a second estimate using the ratio of O_3 and NO_Z sensitivities (S) to on-road mobile (MOB) and electricity generating unit (EGU) source emissions:

$$OPE^{nodep} = \frac{S_{O_3,MOB} + S_{O_3,EGU}}{S_{NO_Z,MOB} + S_{NO_Z,EGU}} \quad \text{(Eq. 7.2)}$$

Where each sensitivity term (*S*) is in units *ppm*. This equation estimates the ratio of the number of O_3 molecules produced per NO_Z molecule attributable to emissions from two sources. We focus on EGU and mobile sources because they make up most of the anthropogenic NO_X emissions. Equation 7.2, however, has one important omission; it does not account for O_3 and NO_Z produced that deposits out (indeed, none of the empirical or chemical transport model methods described to this point explicitly account for deposition). We rectify this by adding source-specific deposition terms:

$$OPE = \frac{\left(S_{O_3,MOB} + S^{DEP}_{O_3,MOB}\right) + \left(S_{O_3,EGU} + S^{DEP}_{O_3,EGU}\right)}{\left(S_{NO_Z,MOB} + S^{DEP}_{NO_Z,MOB}\right) + \left(S_{NO_Z,EGU} + S^{DEP}_{NO_Z,EGU}\right)} \quad \text{(Eq. 7.3)}$$

Where $S^{dep}_{i,j}$ is the sensitivity of species *i* deposition to emissions from source *j* (again, all sensitivity terms are in units *ppm*). CMAQ-DDM outputs the total (wet and dry) O_3 deposition (e.g., $DEP^*_{O_3}$ for O_3) and the sensitivities to total deposition from mobile ($S^{DEP^*}_{O_3,MOB}$) and EGU ($S^{DEP^*}_{O_3,EGU}$) sources. Deposition sensitivities, however, are in units of mass deposited per area; therefore, they require a conversion to concentration units. We achieve this for each source and species by dividing the total deposition sensitivity (e.g., $S^{DEP^*}_{O_3,MOB}$) by the total deposition (e.g., $DEP^*_{O_3}$) and multiplying by the concentration. For example:

$$S^{DEP}_{O_3,MOB} = \frac{S^{DEP^*}_{O_3,MOB}}{DEP^*_{O_3}} * O_3 \quad \text{(Eq. 7.4)}$$

The CMAQ deposition-corrected sensitivities method presents a further opportunity—estimating source-specific OPEs:

$$OPE_{MOB} = \frac{\left(S_{O_3,MOB} + S^{DEP}_{O_3,MOB}\right)}{\left(S_{NO_Z,MOB} + S^{DEP}_{NO_Z,MOB}\right)} \quad \text{(Eq. 7.5)}$$

$$OPE_{EGU} = \frac{\left(S_{O_3,EGU} + S^{DEP}_{O_3,EGU}\right)}{\left(S_{NO_Z,EGU} + S^{DEP}_{NO_Z,EGU}\right)} \quad \text{(Eq. 7.6)}$$

7.4 Results and Discussion

7.4.1 Observation-based OPE

We plotted mean 2-3 p.m. O_3 against mean 2-3 p.m. NO_Z for days with PS* greater than the 80th percentile. The resulting relationships have positive slopes and negative second derivatives, corroborating results from previous studies (Daum et al. 1996; Lin, Trainer, and Liu 1988) (Figure B-1, Figure B-2 for results obtained selecting wider PS* range). NO_X and NO_Z levels are typically lower in the later years and are driven by reductions in NO_X emissions (Hidy and Blanchard 2015; Henneman et al. 2017) (Figure B-3). OPE results from linear and log models are presented in the supplemental (Figure B-4); here we focus on the cubic spline model results.

Intercepts at each of the sites represent an estimate of background concentrations, i.e., expected concentrations when NO_Z levels are quite small due to both low local NO_X

emissions and/or complete deposition of NO_Z. These intercepts likely represent an upper bound of the background for two reasons. First, the spline model's restriction to linearity outside the outermost knots is rather stringent, and may miss small slope increases at the lowest NO_Z concentrations. Second, differences between O_3 and NO_Z deposition are magnified at low concentrations; accounting for this would shift the points plotted in Figure B-1 rightward, thus creating a lower intercept.

Background O_3 varies for each site with meteorological conditions, and long-term climate and emissions changes. Using a Monte Carlo sampling technique, we fit multiple spline models at each site using fixed intercepts selected from the normal distribution around the mean estimated in the original model fit, and calculated an uncertainty distribution of the slopes at each NO_Z concentration (Figure 7-2). This method was chosen because of the high uncertainty in background O_3 and OPE at low NO_Z concentrations and the difficulty in assessing individual model parameter error in spline models. Uncertainty in OPE is larger at sites closer to the Gulf of Mexico.

All O_3-NO_Z spline model intercepts fall between 30.2 (GFP^U) and 47.2 (OAK^R) (Figure 7-1). Urban sites tend to have lower intercepts than rural sites, and the two sites closest to the Gulf of Mexico (GFP^U & PNS^U) have the lowest intercepts with the greatest uncertainty. CTR^R, in Alabama, stands out as the rural site with the lowest intercept. Background estimates align with previous estimates of background estimates in the southeast (Lefohn et al. 2014).

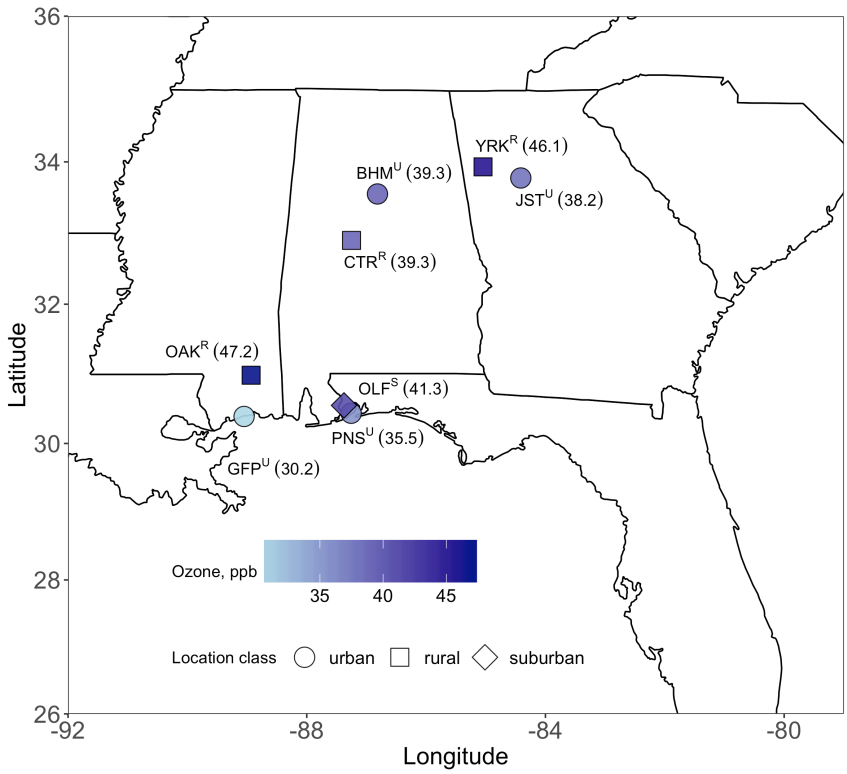


Figure 7-1: Map of SEARCH sites, colored by ozone intercept measured by the spline fit methods and shaped by location class.

Empirical OPEs at JST^U and YRK^R, two sites in Georgia, have similar shapes (Figure 7-2), and OPEs at the rural site are slightly higher than OPEs at the urban site. This trend repeats at BHM^U and CTR^R, with the rural site (CTR^R) exhibiting greater OPEs across all NO_Z levels. The remaining sites (GFP^U, OAK^R, PNS^U, OLF^S) have OPEs that differ from the urban-rural relationships at the four inland sites. Less data available for these sites contributes to larger uncertainty bounds. GFP^U, OLF^S and PNS^U are near the Gulf of Mexico; the air from the Gulf causes OPEs at these sites to increase dramatically at low NO_Z because ozone is efficiently transported over the water with relatively little deposition—due, in part, to its low solubility—while nitric acid is very soluble and will also react with sea salt, leading to larger particles that readily deposit. Neuman et al. (2009) suggest high variability in background depending on wind direction in Houston; these results reflect large uncertainties in coastal sites compared to inland found in the present study (Neuman et al. 2009). OLF^S, a suburban site somewhat inland from PNS^U, has a higher OPE at elevated NO_Z levels and lower OPE at low NO_Z levels than PNS^U. The two

sites nearest the Gulf of Mexico (GFP^U & PNS^U) have the lowest intercepts and the highest OPEs at low NO_Z.

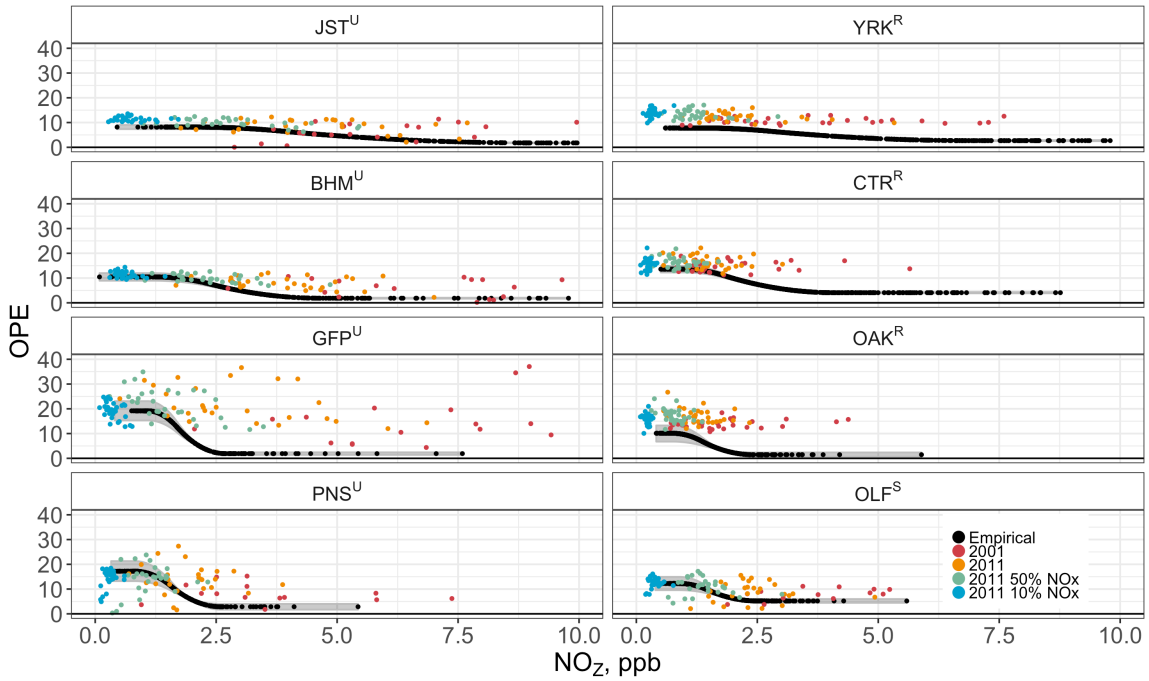


Figure 7-2: OPE at each SEARCH site estimated by the empirical spline and CMAQ deposition-corrected sensitivities approaches. Empirical values represent all days with PS > 80th percentile, and CMAQ values are days in July. Grey shading is the 95% confidence interval around the spline model.*

7.4.2 Chemical Transport Model-based OPE

OPE^{nodep} and OPE^{BF} have a slope near one and R^2 of 0.41 and 0.27, respectively, for two July 2011 comparisons: the change from base emissions to 50% NO_X emissions and from 50% NO_X emissions to 10% NO_X emissions (Figure B-5). The comparison controls for meteorology, but neither controls for deposition; still, the results lend credibility to the sensitivities approach. Without correcting for deposition, over-land average July 2001 and 2011 OPE values were 11.0 and 13.8, a small difference compared to the with-deposition values (Table 7-1). Remaining discussions focus on deposition-corrected OPEs (Equations 7.3, 7.5, and 7.6); all summary statistics are reported for the portion of the domain encompassed by the superimposed polygon in Figure 7-3 to exclude boundary effects and impacts of high OPEs over water.

Table 7-1: Information on the modeling runs. Emissions are summed over the entire domain, while concentrations and OPEs are averaged over the polygon superimposed on each of the plots in Figure 7-3.

	July NO_X Emissions, 10^3 tons			Mean July Concentrations, ppb (std. dev.)			Mean July OPE (std. dev.)			
	ALL	MOB	EGU	O_3	NO_Z	NO_X	OPE	OPE^{nodep}	OPE_{MOB}	OPE_{EGU}
2001	1,364	773	432	54.7 (4.7)	2.6 (0.9)	1.6 (2.7)	11.2 (3.1)	11.0 (3.3)	10.8 (4.6)	12.8 (3.0)
2011	679	364	169	48.1 (5.5)	1.8 (0.6)	1.1 (1.3)	14.0 (3.3)	13.8 (3.3)	13.8 (3.5)	15.3 (3.6)
2011 50% NOx	339	182	85	39.5 (4.7)	1.1 (0.4)	0.5 (0.6)	15.7 (2.3)	15.7 (2.4)		
2011 10% NOx	68	36	17	28.6 (3.6)	0.3 (0.1)	0.1 (0.1)	16.4 (1.7)	16.4 (1.7)		

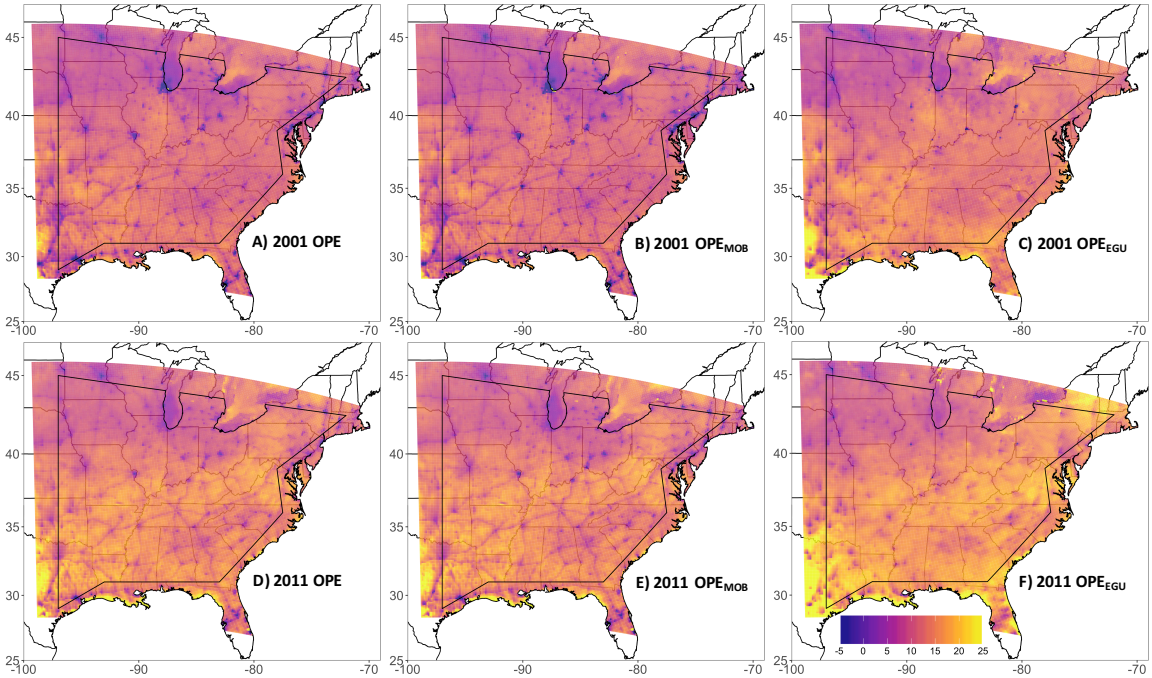


Figure 7-3: CMAQ-modeled mean daily July deposition-corrected OPE (2-3 p.m.) values for 2001 (A-C) and 2011 (D-F), including total OPE (A&D), mobile OPE (B&E), and EGU OPE (C&F).

CMAQ simulations in July yielded consistently higher OPEs in rural areas than urban areas, and generally decreasing OPE in the north of the domain than the south (Figure 7-3). Between 2001 and 2011, average over-land OPEs increased across the domain from 11.2 to 14.0 (Table 7-1) and at each of the SEARCH locations (Figure B-6, Table B-1). Between 2001 and 2011, total NO_X emissions decreased 50%, and OPE increased 25%; a subsequent 50% decrease in emissions would lead to only a 15% increase in OPE; a further 80% emissions cut would increase OPE 4% (Figure B-7). Average July overland OPEs do not exceed 25, even at the most extreme NO_X emission reduction (Figure B-8). Even on

July days with the 2nd-highest simulated O_3 concentrations in each location, OPE does not exceed 27. These results corroborate the empirical results that show that OPE increases to location-specific upper limits at lower NO_X.

Mobile and EGU OPEs differ substantially across the domain. Mobile OPEs are slightly lower than the total OPEs in 2001 and 2011 (Table 7-1), and their spatial pattern generally mirrors that of the total (Figure 7-3). EGU OPEs are generally greater than the total OPEs, with low spots centered near individual plants, which creates a pattern with high standard deviation.

CMAQ OPE results agree with empirical OPEs at low NO_Z concentrations for most sites, although CMAQ OPE estimates at rural sites are slightly higher (Figure 7-2). Note that in this comparison, empirical results use days across the entire period with PS* > 80th percentile, and CMAQ values use only July days in 2001 and 2011. Further, the three coastal sites (GFP^U, OLF^S and PNS^U) are not located within the over-land domain used for analysis (Table 7-1), and direct comparisons between the methods are inherently problematic because both are models with uncertainties and bias.

7.4.3 Implications

Rising OPEs with decreasing NO_Z concentrations suggests a NO_X-limited regime, with increasing effectiveness at reducing O_3 concentrations for each subsequent avoided NO_X molecule emitted. We show this for the empirical models by integrating under the OPE curves in Figure 7-2. For a 0.5 ppb NO_Z concentration decrease, the expected change in O_3 concentrations (ΔO_3) increases with decreasing starting NO_Z concentrations for all sites (Figure 7-4). That ΔO_3 reaches a maximum at the lowest NO_Z concentrations is somewhat a function of the choice of spline model; however, CMAQ results provide evidence for a maximum OPE (and subsequently a maximum ΔO_3). 2011 OPEs across the domain and at all SEARCH sites decreased at a decreasing rate with decreasing NO_X emissions (Table B-1); at CTR^R, GFP^U, and OAK^R, the reduction from 50% to 10% NO_X emissions decreased OPE.

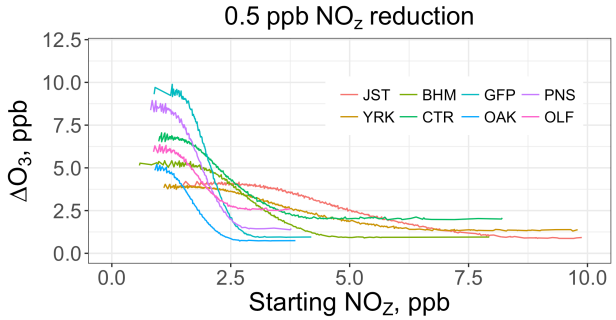


Figure 7-4: Change in O_3 expected from a 0.5 ppb reduction in NO_Z from the value observed at each starting NO_Z for days in July. Starting NO_Z were taken as the daily NO_Z values at each site on days with $PS^* > 80^{th}$ percentile.

We tested the ability of the empirical spline model to reproduce CMAQ-modeled changes in O_3 for a 50% NO_X emissions reduction at the SEARCH sites (Figure B-9). For *relative* reductions (as opposed to *absolute* reductions addressed above), both models find decreasing effectiveness at lower starting NO_Z concentrations (since relative reductions refer to decreasing emissions reductions). The models generally agree for all sites. The comparison assumes that NO_X emissions reductions directly correlate with NO_Z concentration reductions, which is a close (though not perfect) approximation; NO_Z concentrations are linear with reductions using consistent meteorology, but are nonlinear between 2001 and 2011 (Table 7-1, Figure B-7).

CMAQ results yield strong OPE differences between rural and urban areas, and a lesser—though still significant—trend with decreasing latitude (Figure B-10). We related average July OPE over the domain to latitude, NO_X, NO_Z, formaldehyde (HCHO—a reaction product of volatile organics), and the organic fraction of NO_Z ($f_{N,org}$) in multiple linear regressions for 2001 and 2011 (Table B-2). While all regression parameters were significant in each iteration of the linear model, only the model with latitude and $f_{N,org}$ yielded a small influence of latitude on OPE. This held for both 2001 and 2011 and reduced the intercept at 29º latitude. Higher OPEs in the south are likely due to increased organics; however, $f_{N,org}$ serves as a sufficient marker for the broader spatial trend because it is spatially consistent even in cities, whereas the other variables (NO_X, NO_Z, and HCHO) are elevated in urban areas (Figure B-11).

Empirical and CMAQ results provide evidence that maximum OPEs are site-specific. At the lowest NO_Z concentrations, the potential for reducing O_3 concentrations through NO_X emissions reductions levels off; however, O_3 concentrations have a positive

relationship with NO_Z at all NO_Z concentrations. This suggests that continuing emissions controls will reduce O_3 concentrations at increasing levels, but the effect will level off at location-specific NO_Z concentrations.

EGU OPEs are on average greater than mobile OPEs across the domain. This suggests greater expected O_3 reductions per absolute NO_X emissions reduction from EGUs than mobile. However, this comparison does not consider the amount of emissions left to reduce—mobile sources emitted more than 2X NO_X than EGU sources in 2011. Since EGU OPE is 33% larger than mobile OPE (15.6 vs. 11.7), more O_3 reduction is available through proportional mobile emission reductions than EGU.

7.4.4 Approach limitations

Various assumptions are inherent in the methods discussed in this study. The empirical method effectively uses a *Eulerian* framework, i.e., stationary measurements of plumes with different histories and degrees of photochemical activity. The approach is subject to varying background pollutant levels, nighttime nitrate chemistry, air parcel transport, varying NO_3 and O_3 deposition rates, and clouds (Ryerson et al. 1998). A *Lagrangian* approach, i.e., when observations are made while traveling with a specific air mass, does not suffer from having air parcels with varying histories, though does suffer from the other issues (e.g., varying deposition rates) and is more difficult to conduct, though has been undertaken using aircraft, particularly following plumes of elevated NO_Y (Ryerson et al. 1998; Sillman 2000; Kleinman et al. 2002; Imhoff et al. 1995; Neuman et al. 2009). The Lagrangian approach typically has led to lower OPEs than those found using Eulerian approaches, although this may be due to higher NO_Z concentrations in plumes, a more limited time to react or because the typical approach is to correlate O_3 with NO_Z, instead of taking the nonlinearity of the system into account (Imhoff et al. 1995; Trainer et al. 1995; Ridley et al. 1998; Thornton et al. 2002).

The empirical approach is further limited because it does not control for changing emissions of other species involved in ozone formation (e.g., VOCs and CO) or long term changes in the meteorology leading to high ozone days (e.g., due to climate change). Sensitivity studies find that O_3 is primarily sensitive to NOx in the Southeast on high ozone days (Blanchard, Hidy, and Tanenbaum 2014; Mazzuca et al. 2016), though results in downtown areas suggest negative sensitivity to local NO_X emissions and VOC sensitivity

even on the highest days (Blanchard, Hidy, and Tanenbaum 2010; Mazzuca et al. 2016; Simon et al. 2014). Urban NO emissions inhibit O_3 production, albeit primarily on days with low photochemical activity (Seinfeld and Pandis 2006). Based on relationships developed above, future urban NO_X emissions will reduce O_3 on high-photochemistry days; however, emissions reductions will lead to continuing increases on low-photochemistry days (primarily in the wintertime).

Recent work has shown that a portion of NO_X in the Southeast may react with organics and hydrolyze to form organic nitrogen aerosol. Fisher et al. (2016) found that 21% of NO_X was lost to $RONO_2$, and that this fraction increases with decreasing NO_X emissions (Fisher et al. 2016). 59% of $RONO_2$ was hydrolyzed to aerosol in the southeastern US in August-September 2013. This portion of NO_Z is not present in NO_Y measurements, and would lead to increasing bias in the empirical OPE model at low NO_Z concentrations.

The empirical and CMAQ-based methods applied here were designed to reduce biases associated with these limitations. In the empirical method, the use of meteorologically similar days with high photochemical activity (PS*) across multi-year datasets limits (but likely does not remove) effects of plumes with markedly different origins. We assessed uncertainty based on a range of assumed background O_3 levels, and found a greater impact on OPE at coastal sites. Using CMAQ-DDM, we controlled for air parcel transport and differing O_3 and NO_Z deposition rates by applying source-specific concentration and deposition sensitivity ratios instead of just concentrations. A limitation of the CMAQ approach is the use of concentrations, deposition, and sensitivities to anthropogenic emissions at a single grid cell, whereas each of these would be impacted by transport and deposition between the source and the grid cell of interest. However, the deposition fields are relatively smooth and spatially similar to the concentration field, suggesting this issue is of limited importance (Figure B-11 and Figure B-12).

This study shows that future NO_X emissions reductions remain a viable option for reducing the highest O_3 concentrations. OPE increases with decreasing NO_Z concentrations, but appears to level out or possibly reach a maximum, which implies that absolute NO_X emissions reductions will become more effective at reducing O_3 up to a point, after which the effect will level off. Using the empirical modeling approach, in rural

areas, this maximum OPE has already been attained, and is approximately the case using the chemical transport modeling approach as well.

7.5 Acknowledgements

This material is based upon work supported by Health Effects Institute and the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1148903. We are grateful for discussions with Charlie Blanchard regarding measurements made at SEARCH network sites.

7.6 References

Blanchard, C. L., G. M. Hidy, and S. Tanenbaum. 2014. "Ozone in the Southeastern United States: An Observation-Based Model Using Measurements from the SEARCH Network." *Atmospheric Environment* 88. Elsevier Ltd: 192–200. doi:10.1016/j.atmosenv.2014.02.006.

Blanchard, C.L., G.M. Hidy, and S. Tanenbaum. 2010. "NMOC, Ozone, and Organic Aerosol in the Southeastern United States, 1999–2007: 2. Ozone Trends and Sensitivity to NMOC Emissions in Atlanta, Georgia." *Atmospheric Environment* 44 (38). Elsevier Ltd: 4840–49. doi:10.1016/j.atmosenv.2010.07.030.

Byun, Daewon, and Kenneth L. Schere. 2006. "Review of the Governing Equations, Computational Algorithms, and Other Components of the Models-3 Community Multiscale Air Quality (CMAQ) Modeling System." *Applied Mechanics Reviews* 59 (2): 51. doi:10.1115/1.2128636.

Camalier, Louise, William Cox, and Pat Dolwick. 2007. "The Effects of Meteorology on Ozone in Urban Areas and Their Use in Assessing Ozone Trends." *Atmospheric Environment* 41 (33): 7127–37. doi:10.1016/j.atmosenv.2007.04.061.

Daum, P. H., L. I. Kleinman, L. Newman, W. T. Luke, J. Weinstein-Lloyd, C. M. Berkowitz, and K. M. Busness. 1996. "Chemical and Physical Properties of Plumes of Anthropogenic Pollutants Transported over the North Atlantic during the North Atlantic Regional Experiment." *Journal of Geophysical Research* 101 (D22): 29029. doi:10.1029/95JD03163.

Fisher, Jenny A., Daniel J. Jacob, Katherine R. Travis, Patrick S. Kim, Eloise A. Marais, Christopher Chan Miller, Karen Yu, et al. 2016. "Organic Nitrate Chemistry and Its Implications for Nitrogen Budgets in an Isoprene- and Monoterpene-Rich Atmosphere: Constraints from Aircraft (SEAC4RS) and Ground-Based (SOAS) Observations in the Southeast US." *Atmospheric Chemistry and Physics* 16 (9): 5969–91. doi:10.5194/acp-16-5969-2016.

Godowitch, James M., Robert C. Gilliam, and S. Trivikrama Rao. 2011. "Diagnostic Evaluation of Ozone Production and Horizontal Transport in a Regional Photochemical Air Quality Modeling System." *Atmospheric Environment* 45 (24). Elsevier Ltd: 3977–87. doi:10.1016/j.atmosenv.2011.04.062.

Hansen, D. Alan, Eric S. Edgerton, Benjamin E. Hartsell, John J. Jansen, Navaneethakrishnan Kandasamy, George M. Hidy, and Charles L. Blanchard. 2003. "The Southeastern Aerosol Research and Characterization Study: Part 1—Overview." *Journal of the Air & Waste Management Association* 53 (12): 1460–71. doi:10.1080/10473289.2003.10466318.

Hansen, D Alan, Eric Edgerton, Ben Hartsell, John Jansen, Harriet Burge, Petros Koutrakis, Christine Rogers, et al. 2006. "Air Quality Measurements for the Aerosol Research and Inhalation Epidemiology Study." *Journal of the Air & Waste Management Association (1995)* 56 (10): 1445–58. doi:10.1080/10473289.2006.10464549.

Hastie, Trevor, Robert Tibshirani, Jerome Friedman, and James Franklin. 2009. *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. 2nd ed. Stanford, CA: Springer. doi:10.1007/BF02985802.

He, H., J. W. Stehr, J. C. Hains, D. J. Krask, B. G. Doddridge, K. Y. Vinnikov, T. P. Canty, et al. 2013. "Trends in Emissions and Concentrations of Air Pollutants in the Lower Troposphere in the Baltimore/Washington Airshed from 1997 to 2011." *Atmospheric Chemistry and Physics* 13 (15): 7859–74. doi:10.5194/acp-13-7859-2013.

Henneman, Lucas RF, Howard H Chang, David Lavoue, James A Mulholland, and Armistead G Russell. 2017. "Accountability Assessment of Regulatory Impacts on Ozone and PM2.5 Concentrations Using Statistical and Deterministic Pollutant Sensitivities." *Air Quality Atmosphere and Health*. doi:10.1007/s11869-017-0463-2.

Henneman, Lucas RF, Heather A Holmes, James A Mulholland, and Armistead G Russell. 2015. "Meteorological Detrending of Primary and Secondary Pollutant Concentrations: Method Application and Evaluation Using Long-Term (2000-2012) Data in Atlanta." *Atmospheric Environment* 119. Elsevier Ltd: 201–10. doi:10.1016/j.atmosenv.2015.08.007.

Henneman, Lucas RF, Cong Liu, Yongtao Hu, James A Mulholland, and Armistead G Russell. n.d. "Air Quality Modeling for Accountability Research: Operational, Dynamic, and Diagnostic Evaluation."

Hidy, George M, and Charles L Blanchard. 2015. "Precursor Reductions and Ground-Level Ozone in the Continental United States." *Journal of the Air & Waste Management Association (1995)* 65 (10): 1261–82. doi:10.1080/10962247.2015.1079564.

Imhoff, R E, R Valente, J F Meagher, and M Luria. 1995. "The Production of O-3 in an Urban Plume: Airborne Sampling of the Atlanta Urban Plume." *Atmospheric Environment* 29 (17): 2349–58.

Kleinman, L I, P H Daum, Y N Lee, L J Nunnermacker, S R Springston, J Weinstein-Lloyd, and J Rudolph. 2002. "Ozone Production Efficiency in an Urban Area." *Journal of Geophysical Research-Atmospheres* 107 (D23): 4733. doi:Doi 10.1029/2002jd002529.

Lefohn, Allen S., Christopher Emery, Douglas Shadwick, Heini Wernli, Jeagun Jung, and Samuel J. Oltmans. 2014. "Estimates of Background Surface Ozone Concentrations in the United States Based on Model-Derived Source Apportionment." *Atmospheric Environment* 84. Elsevier Ltd: 275–88. doi:10.1016/j.atmosenv.2013.11.033.

Lin, X., M. Trainer, and S. C. Liu. 1988. "On the Nonlinearity of the Tropospheric Ozone Production." *Journal of Geophysical Research* 93 (D12): 15879. doi:10.1029/JD093iD12p15879.

Liu, S.C., M. Trainer, F. C. Fehsenfeld, D. D. Parrish, E. J. Williams, D. W. Fahey, G. Hobler, and P. C. Murphy. 1987. "Ozone Production in the Rural Troposphere and the Implications for Regional and Global Ozone Distributions." *Journal of Geophysical Research* 92 (D4): 4191–4207.

Mazzuca, Gina M., Xinrong Ren, Christopher P. Loughner, Mark Estes, James H. Crawford, Kenneth E. Pickering, Andrew J. Weinheimer, and Russell R. Dickerson. 2016. "Ozone Production and Its Sensitivity to NO*x* and VOCs: Results from the DISCOVER-AQ Field Experiment, Houston 2013." *Atmospheric Chemistry and Physics* 16 (22): 14463–74. doi:10.5194/acp-16-14463-2016.

Neuman, J. a., J. B. Nowak, W. Zheng, F. Flocke, T. B. Ryerson, M. Trainer, J. S. Holloway, et al. 2009. "Relationship between Photochemical Ozone Production and NOx Oxidation in Houston, Texas." *Journal of Geophysical Research: Atmospheres* 114 (10): 1–12. doi:10.1029/2008JD011688.

R Development Core Team. 2012. "R: A Language and Environment for Statistical Computing." *R Foundation for Statistical Computing Vienna Austria*. Vienna, Austria. doi:ISBN 3-900051-07-0.

Ridley, Brian A., James G. Walega, J.-F. Lamarque, F.E. Grahek, M. Trainer, G. Hubler, X. Lin, and Fredrick C Fehsenfeld. 1998. "Measurements of Reactive Nitrogen and Ozone to 5-Km Altitude in June 1990 over the Southeastern United States." *Journal of Geophysical Research* 103 (D7): 8369–88.

Ryerson, T. B., M. P. Buhr, G. J. Frost, P. D. Goldan, J. S. Holloway, G. Hübler, B. T. Jobson, et al. 1998. "Emissions Lifetimes and Ozone Formation in Power Plant Plumes." *Journal of Geophysical Research* 103 (D17): 22569. doi:10.1029/98JD01620.

Seinfeld, John H, and Spyros N Pandis. 2006. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. *Atmospheric Chemistry and Physics*. Vol. 2nd. http://www.knovel.com/knovel2/Toc.jsp?BookID=2126.

Shen, L., L. J. Mickley, and A. P K Tai. 2015. "Influence of Synoptic Patterns on Surface Ozone Variability over the Eastern United States from 1980 to 2012." *Atmospheric Chemistry and Physics* 15 (19): 10925–38. doi:10.5194/acp-15-10925-2015.

Shen, Lu, and Loretta J Mickley. 2017. "Seasonal Prediction of US Summertime Ozone Using Statistical Analysis of Large Scale Climate Patterns." *PNAS* 114 (10): 2491–96. doi:10.1073/pnas.1610708114.

Sillman, S. 2000. "Ozone Production Efficiency and Loss of NOx in Power Plant Plumes: Photochemical Model and Interpretations of Measurements in Tennessee." *Jounal of Geophysical Research*.

Simon, Heather, Adam Reff, Benjamin Wells, Jia Xing, and Neil Frank. 2014. "Ozone Trends across the United States over a Period of Decreasing NOx and VOC Emissions." *Environmental Science & Technology* 49 (1): 186–95. doi:10.1021/es504514z.

Thornton, J. A., P. J. Wooldridge, R. C. Cohen, M Martinez, H Harder, W. H. Brune, E.J. Williams, et al. 2002. "Ozone Production Rates as a Function of NOx Abundances and HOx Production Rates in the Nashville Urban Plume." *Journal of Geophysical Research* 107 (D12): 1–17. doi:Artn 4146\rDoi 10.1029/2001jd000932.

Trainer, M, D D Parrish, M P Buhr, R B Norton, F C Fehsenfeld, K G Anlauf, J W Bottenheim, et al. 1993. "Correlation of Ozone with NOy in Photochemically Aged Air." *Journal of Geophysical Research: Atmospheres* 98 (D2): 2917–25. doi:10.1029/92JD01910.

Trainer, M, B a Ridley, M P Buhr, G Kok, J Walega, G Hubler, D D Parrish, and F C Fehsenfeld. 1995. "Regional Ozone and Urban Plumes in the Southeastern United States - Birmingham, a Case Study." *Journal of Geophysical Research-Atmospheres* 100 (D9): 18823–34. doi:10.1029/95JD01641.

U.S. EPA. 2015. *National Ambient Air Quality Standards for Ozone Final Rule*. *40 CFR Parts 50,51,52,53 and 58*. Vol. 80.

Vijayaraghavan, Krish, Allison DenBleyker, Lan Ma, Chris Lindhjem, and Greg Yarwood. 2014. "Trends in on-Road Vehicle Emissions and Ambient Air Quality in Atlanta, Georgia, USA, from the Late 1990s through 2009." *Journal of the Air & Waste Management Association* 64 (7): 808–16. doi:10.1080/10962247.2014.892039.

Xing, J., J. Pleim, R. Mathur, G. Pouliot, C. Hogrefe, C. M. Gan, and C. Wei. 2013. "Historical Gaseous and Primary Aerosol Emissions in the United States from 1990 to 2010." *Atmospheric Chemistry and Physics* 13 (15): 7531–49. doi:10.5194/acp-13-7531-2013.

Zurbenko, Igor G. 1991. "Spectral Analysis of Nonstationary Time Series." *International Statistical Review* 39: 163–73.

CHAPTER 8. ASSESSING EMISSIONS LEVELS AND COSTS ASSOCIATED WITH CLIMATE AND AIR POLLUTION POLICIES IN SOUTH AFRICA[6]

8.1 Abstract

Affordable energy supply and reductions in emissions of local air pollution and greenhouse gases are each important aspects of South Africa's goals. Many traditional solutions, however, work in contradiction to one another. This work investigates effects on estimated emissions and costs of mitigation strategies using the Greenhouse Gas and Air Pollution Interaction Synergies (GAINS) model to identify policies that satisfy multiple goals. Eight scenarios that describe air pollution control options and mixes of energy production technologies are implemented in GAINS, which quantifies country-wide air pollution and greenhouse emissions and costs of controls. Emissions and costs trajectories are compared to the business as usual case, which projects CO_2 emissions to increase by 60% by 2050 compared to 2015. Results show that replacing all coal generation with renewables reduces CO_2 emissions in 2050 by 8% compared to 2015, and that aggressive policy targeting the whole energy sector reduces CO_2 emissions in 2050 by 40%. GAINS is used to show co-benefits and tradeoffs of each scenario, such as reductions in emissions control costs that accompany a switch to renewables. The approach provides supporting evidence for policies that exploit co-benefits and avoid contradictions by assessing multiple aspects of the energy sector within the integrated framework provided by the GAINS modeling platform.

8.2 Introduction

South Africa is committed to growing their economy. A key component of this goal is to provide and expand access to affordable energy to all its citizens and to industries.

[6] A modified version of this chapter has been published as: Lucas R.F. Henneman, Peter Rafaj, Harold J. Annegarn, Carmen Klausbruckner, Assessing emissions levels and costs associated with climate and air pollution policies in South Africa, Energy Policy, Volume 89, February 2016, Pages 160-170, ISSN 0301-4215, https://doi.org/10.1016/j.enpol.2015.11.026. Supplementary data related to this article can be found at http://www.sciencedirect.com/science/article/pii/S0301421515302044.

This energy, however, must be produced in a manner that both protects the health of citizens and the environment, and upholds the country's international commitments to address emissions of greenhouse gases (GHGs).

Many air pollutants are emitted when fossil fuels are combusted. Particulate matter (PM) has been linked in a number of studies to negative health effects. Exposure to household air pollution from solid fuels and ambient airborne PM ranked second and eighth in attributable disease burden in the 2013 Global Burden of Disease study (Lim et al., 2013). Other air pollutants, such as sulfur dioxide (SO_2) and oxides of nitrogen (NO_X = $NO + NO_2$), contribute to formation of secondary PM and ozone (O_3), a second air pollutant with known health effects, respectively.

Besides air pollutants that affect local human and ecological health, South Africa is concerned with reducing its contribution to rising GHGs concentrations in the atmosphere. Fossil fuel use in South Africa contributes to the country's standing as the world's 13th largest emitter of GHGs (US EIA, 2013).

The complex regulatory frameworks of energy, air pollution, and GHG policies are all related, and it is necessary to investigate them under a common framework. An integrated approach allows for co-benefits to be identified and exploited and for contradictions to be avoided. The Greenhouse Gas and Air Pollution Interactions and Synergies (GAINS) model provides such a framework (IIASA, 2012). This study presents an impact analysis of policies targeting the energy sector in South Africa, and investigates potentials for air pollutant and GHG emissions reductions under various pollution control and energy activity scenarios. To maintain coherence with the continuing national discussion on mitigation of climate change, the selected scenarios have been styled on commonly used scenarios in order to simulate policy options, such as those presented in the report 'Study to Examine the Potential Socio-Economic Impact of Measures to Reduce Air Pollution from Combustion' ((Scorgie et al., 2003) and the Long Term Mitigation Strategies (LTMS) report, a policy document that describes options available to South Africa to curb its carbon dioxide (CO_2) emissions (ERC, 2007). Emissions of SO_2, particulates with diameter less than 2.5 μm ($PM_{2.5}$), NO_X, and CO_2 are estimated for each scenario along with the cost of air pollution controls.

The paper begins with a review of relevant energy policy in South Africa. A companion paper provides a more detailed analysis of the policy and legal framework in South Africa (Klausbruckner et al., 2016). A description of the GAINS model precedes a detailing of the scenarios modeled and their relevance to policy options that South Africa has. Results from the GAINS model for each of these scenarios are summarized in the Results section and analyzed further in the Discussion section. The results will inform policy makers and future studies on impacts of energy systems both in South Africa and other countries that wish to apply the GAINS model for an integrated analysis to shape policy.

8.3 Policy framework

8.3.1 Energy policy in South Africa

South Africa has an energy-intensive economy. The country's index of primary energy supply per US dollar of gross domestic product (GDP) based on purchasing power parity is 11.7 mega joules (MJ), which outranked both Asia (7.9 MJ/USD) and Latin America (6.7 MJ/USD) in 2007 (Winkler, 2007). Over 90% of the electricity in South Africa is produced by burning coal (Edkins et al., 2010). Industry, transport, mining, and agriculture make up the top four demand sectors (Winkler, 2007).

Historically, South Africa had an excess of electricity generation capacity to supply the demand, which contributed to the country having among the lowest electricity prices in the world ($0.02/kWh or less). For diverse reasons, including a protracted period of 20 years in which no new generating plants were built, this situation has changed in recent years (2008 onwards) so that supply has not always been able to meet demand. From 2008 onwards, the country saw increasing prices and intermittent load shedding (scheduled blackouts on a planned regional basis) (Edkins et al., 2010; Tait & Winkler, 2012). In 2015, Eskom, the state-owned electricity utility, planned to begin operation of the first of two new coal-fired power plants currently under construction, each with a planned capacity of 4,800 Megawatt electric (MWe) (Kiratu, 2010).

In 2007, the South African Government commissioned the LTMS report, to investigate pathways that the country could take to mitigate national emissions of GHGs (ERC, 2007). This report presented scenarios and policy options that would allow the country to follow a "required by science" emissions pathway that aligns with worldwide

emissions reductions of 30-40% below 2003 levels in 2050 necessary to keep global warming below 2ºC (IPCC, 2001). Four policy timeline options are proposed as necessary for achieving the required by science pathway in the report– "Start Now," "Scale Up," "Use the Market," and "Reach for the Goal". Each of these describes a suite of policies that, if each implemented in full, would achieve 64% of the reductions in GHG emissions needed in the 'required by science' scenario. The options suggest a combination of investment in both positive and negative cost technologies (i.e., technologies that do not repay their value over time), taxes and incentives, and behavioral change. This document has formed the basis for national climate change policy since 2007, however, there has been a lack of coordinated effort to implement these policies. The LTMS, for instance, proposed eliminating the commissioning of new coal plants, and Eskom has since begun building the two new plants mentioned above (Kiratu, 2010), while the low-carbon (but otherwise still controversial) nuclear power option has been subjected to repeated delays, although a tentative deal was reached in 2014 with the Russian company Rosatom to supply 9.6 GWe of nuclear capacity to South Africa by 2030 (WNA, 2014).

8.3.2 Air quality policy in South Africa

South Africa routinely experiences levels of air pollution that are detrimental to human health in many areas of the country. Besides ambient air quality issues caused by industrial and mobile sources, many poorer communities suffer negative health outcomes due to exposure to high levels of pollution from fossil fuels used for cooking, heating, and lighting (Scorgie et al., 2003; Pauw et al., 2008).

The National Environmental Management Air Quality Act, promulgated in 2004, formed the legal basis for defining Minimum Emissions Standards (MES) for regulating gaseous and particulate emissions from industrial operations. The MES apply to both new and existing plants in industrial and electricity generating sectors, were amended in 2013 and come into effect in 2015. The effectiveness of this regulation may be limited, however, as many of the major emitters have applied for deferments of the date of compliance with the MES (Myllyvirta, 2014). Extensions through 2020 have been granted to certain plants for SO_2 and NO_X emissions.

Regarding mobile emissions, the sale of fuel with high sulfur content (up to 500 ppb) has limited the ability of automobile manufacturers to sell vehicles with engines that

employ the current cleanest technology. In 2012, the government issued a Government Notice specifying the compulsory introduction of fuels that meet EURO 5 standards (less than 10 ppm sulfur) by July 2017. While there has been some effort to begin marketing these fuels already, the enforcement of regulations on low-sulfur fuels may be postponed, which would delay widespread introduction of higher-standard fuels (SAPIA, 2014). A recent agreement by South Africa suggests that enforcement of low-sulfur fuels will occur begin in 2020 (Workshop, 2020).

8.3.3 Climate change mitigation policy in South Africa

In 2009, at the United Nations Framework Convention on Climate Change (UNFCCC) Conference of the Parties (COP 15) in Copenhagen, South African President Zuma pledged a target of CO_2 emissions reductions below 'business as usual' of 34% by 2020 and 42% by 2025 (Kiratu, 2010) (subject to conditions on the provision of financial support). These targets came as a surprise to many South African policy makers and industry leaders at the time (Kiratu, 2010), but there has since been some progress in formulating policies to reduce GHG emissions in line with these pledges. In 2010, the government developed an Integrated Resources Plan (IRP), which set a goal that renewable energy will make up 14% of the electricity generating mix by 2030 (Merven et al., 2014; DOE, 2013). The LTMS report found that a tax on carbon emissions would be the most efficient policy to reduce CO_2 emissions, and such a tax was proposed to begin in 2015, but implementation has been delayed to 2016 (Merven et al., 2014). Results of the LTMS were further used to inform the creation of the 2011 National Climate Change Response White Paper (S.A. Government, 2011). This document laid out plans to peak emissions between 2020 and 2025, plateau until 2035, and decrease annual emissions thereafter. The White Paper states that the government will prioritize policy actions that address both GHG emissions and negative public health outcomes due to local air pollution exposure. The government hopes to implement a solution that achieves reductions in GHG emissions and maintains its competitiveness on the global level (S.A. Government, 2011).

Under the original IRP published in 2010, South Africa would reduce dependence on coal for electricity production from 90% to 65% in 2030. The share of electricity produced would be 20% from nuclear and to 14% from renewables in 2030. An updated IRP published in 2013 presents decision trees to enable decision-making in a changing

economic climate. Under this plan, the Department of Energy aims to fund projects that will install a generating capacity of 3.7 GW (8.4% of total generating capacity) by renewables by 2015; however, the capacity is somewhat behind this goal (Merven et al., 2014). These policies and the deal to increase nuclear capacity show that the government is aware of the need for climate change mitigation policies and is taking action to curb emissions.

8.4 Methods

8.4.1 The GAINS model

The International Institute for Applied Systems Analysis (IIASA) developed the GAINS techno-economic optimizing model to evaluate effects of air pollution controls and GHG emissions reduction policies on air pollutant emissions, public and environmental health, and abatement costs (IIASA, 2012). GAINS was designed to be easily adapted to different world regions, and has been applied to a number of regions and individual countries around the world, including Europe (Amann et al., 2011; Wagner et al., 2013; Winiwarter, 2005), China (Amann et al., 2008), Pakistan (Purohit et al., 2013) and others. Applications have included studying the costs and benefits of climate-mitigation policies, the joint benefits of climate change mitigation and air pollution control policies, and the health benefits of reducing air pollution emissions (Rafaj et al., 2013).

GAINS has the capability to estimate emissions of both GHGs (including CO_2, CH_4, N_2O, and fluorinated gases) and local air pollutants (including SO_2, NO_x, volatile organic compounds–VOCs, NH_3, and PM). At its core, the model applies a single equation that calculates emissions using data entered by the user (Winiwarter, 2005). In past applications, it has been combined with cost-optimization, chemical transport, and epidemiological models. These three extensions require additional models that are not included in the current study; for this application, only the emissions and cost calculations are applied.

$$E_p = \sum_{j,a,t} A_{j,a} ef_{j,a,p} \left(1 - eff_{t,p}\right) X_{j,a,t} \qquad \text{(Eq. 8.1)}$$

j,a,t,p Subscripts denoting sector, activity, abatement technology, and pollutant, repectively

E_p Emissions of pollutant *p*

A_j Activity in sector j

$ef_{j,a,p}$ Uncontrolled emissions factor in sector j of activity a and pollutant p

$eff_{t,p}$ Reduction efficiency of abatement technology t on pollutant p

$X_{j,a,t}$ Implementation rate of technology t sector j and activity a.

The graphical representation of this equation (Figure 8-1) highlights the two inputs–activity (A in the equation above) and emission controls (X in the equation above)–that are adjusted to create the scenarios. Emissions vectors (ef and eff in the equation above) represent constants that are unique to each activity or abatement technology.

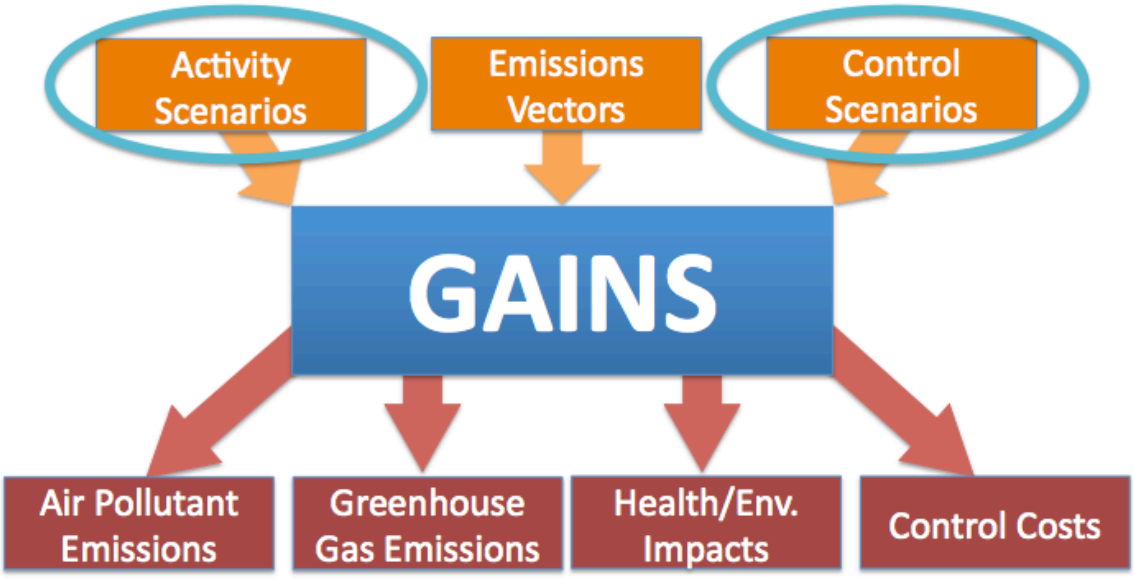


Figure 8-1. Schematic representing the inputs and outputs of the current implementation of GAINS. The ovals denote the inputs that are adjusted to create the eight scenarios assessed in this work.

While GAINS is a powerful tool that can be used to analyze the emissions, public health, and economic outcomes of different scenarios, it does have limitations. First, the activity inputs to GAINS are developed outside of the model. For instance, it does not model electricity demand based on price, which has an effect on the emissions from the electricity generation sector (i.e., GAINS is not an electricity dispatching tool). Other factors, such as population growth, deployment of new power supply technologies (e.g. renewables), and improved efficiency are modeled in separate energy models, and not within GAINS.

Control costs are estimated in GAINS by assuming a cost for each control technology per amount of pollutant controlled. These are multiplied by the amount of pollutant that is controlled with each technology at a given application rate. A 4% interest rate is used in this study.

This paper presents the initial implementation of GAINS in South Africa, in which only a limited range of GAINS outputs are assessed. Other air pollutants (e.g. ammonia, VOCs) and non-CO_2 GHGs, although covered by the GAINS modelling framework, are not included in the analysis presented herein. Functions to model health outcomes require the inclusion of data layers generated by external dispersion and epidemiological models that have not as yet been included in the South African GAINS domain. The current implementation deals with the overall energy and total national emissions of the defined scenarios.

8.4.2 Scenario development

The modeling was performed using eight scenarios (Table 8-1). These were split into two varieties: control scenarios and activity scenarios. Scenarios were designed to reflect similar scenarios presented in previous modeling efforts (for instance, in the LTMS and Gauteng Integrated Energy Strategy (DLGH, 2010; LTMS, 2007).

Table 8-1. Control and Activity scenarios employed in modeling.

Scenario	Acronym	Notes
Baseline Scenario		
Business As Usual	BAU	Based on IEA 6º warming scenario and current legislation
Control Scenarios		
No Further Controls[a]	NFC	Freeze current control levels after 2015
Maximum Feasible Controls[a]	MFC	Invest in best available technologies on all sources by 2030
Activity Scenarios		
Clean fuels in DOMestic sector[a]	DOM	Replace solid fuels with LPG
Clean Coal Technologies[a]	CCT	Construct IGCC power plants
Coordinated global mitigation strategy (2ºC)[b]	2DS	Based on IEA 2º warming scenario(IEA , 2012)
RENewables only in electricity generation[b]	REN	No coal in electricity generation by 2050
RENewables Tradeoff[b]	RENT	Increasing domestic solid fuels with higher electricity prices

[a]Air pollution reduction scenario

[b]Climate change mitigation scenario

8.4.3 Baseline scenario

The *Business As Usual* (BAU) scenario is based on energy projections developed by the International Energy Agency (IEA, 2012). In their 2012 report, the International Energy Agency defines this scenario as the 6º-warming scenario. The BAU scenario

represents the current policy environment in South Africa in terms of emissions standards, and the energy activity mix is grounded in the projected GDP and population growth presented in Figure 8-2a. The energy consumption mix is presented in Figure 8-2b, and is dominated by coal. Current air pollution control levels, adapted from Cofala et al. (2012) and Klimont et al. (2016), imply that future control requirements will be at least as stringent as current standards (supplemental, S1). The inputs assume that there will be compliance of existing regulations, except for a few instances for which there have been submissions requesting a delay of enforcement of the regulations. Two instances in which the delayed implementation are taken into account are the mobile sector implementing Euro 5 fuel standards beyond the 2017 deadline, and formal submissions for delays in complying with minimum emissions standards for PM and SO_2 from existing power plants.

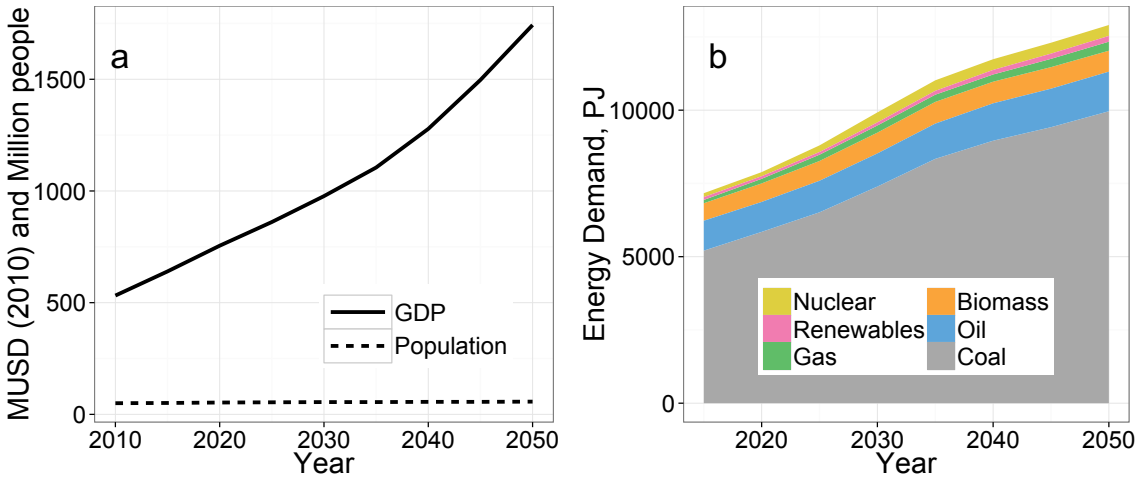


Figure 8-2. Projected GDP (in 2010 USD) and population in South Africa. Estimated GDP in 2015 (2050) is 640 (1740) million USD, and population is 51 (57) million people (IEA, 2012). In South Africa, the growth in GDP (population) of 170% (10%) are lower than the rest of Africa (260% growth in GDP and 90% growth in population). b) Projected energy consumption by fuel in South Africa the BAU scenario.

8.4.4 Control scenarios

Two control strategies are implemented in GAINS (Table 8-1). The *No Further Controls* (NFC) strategy assumes that all emitters are controlled at 2015 levels as implemented in the BAU scenario until 2050. This represents the control scenario with the least costs, since no further investments would be needed after 2015 except to maintain controls. The *Maximum Feasible Controls* (MFC) scenario represents an option in which emitters in all sectors would implement the best available technology to control all

pollutants. These technologies include flue gas desulphurization, high efficiency dedusters, and selective catalytic reduction technologies on industrial sources and power plants, improved stoves and burning techniques in the domestic sector, and the most stringent Euro standards and low-sulfur fuels in the transportation sector (supplemental, S1).

The implementation in GAINS assumes that all MFC controls are installed by 2030. These two scenarios form the likely space that the actual emission path will take, and serve as comparisons for the BAU and activity scenarios.

8.4.5 Activity scenarios

Five activity scenarios are implemented to demonstrate how altering the suite of energy activities affects emissions and investments in air pollution controls (Table 8-1). Activity scenarios are designed as strategies to reduce local air pollution, mitigate GHG emissions, increase energy security through diversity, etc.; however, it is often the case that the benefits of one strategy extend beyond the primary goal. It may be that strategies have unwelcome consequences as well, such as the abandonment of electricity use for cooking due to increasing prices. The GAINS framework allows for the analysis of both multiple benefits and unwelcome consequences and the comparison across scenarios.

Two air pollution reduction strategies are implemented: *Clean fuels in the DOMestic sector* (DOM) and *Clean Coal Technologies* (CCT) in electricity generation. The first (DOM) is a strategy that would replace solid fuels with clean household fuels (e.g. liquid petroleum gas) in the domestic sector by 2020, with the goal of reducing the most important source of human exposure to PM in South Africa (Scorgie et al., 2003). The second air pollution approach (CCT) is to employ clean coals technologies in the electricity generation sector. This scenario replaces all new electricity generating capacity–as the current capacity is retired in line with its projected lifetime–with high-efficiency integrated coal-gasification combined cycle (IGCC) plants. While IGCC offers the most promising option for the implementation of carbon capture and storage (CCS), and the technology has been mentioned as a potential strategy in multiple policy documents (*Long Term Mitigation Scenarios Technical Summary*, 2007; South African Government, 2011), the technology is not included in the modeling of clean coal options performed here. Reported CO_2 emissions reductions are driven instead by efficiency increases in the power sector.

Two climate change strategies are implemented. The first (2DS) is an adoption of the coordinated global mitigation strategy aimed at achieving the 2º warming climate stabilization target as developed by the International Energy Agency (IEA, 2012). The second (REN) is characterized by the replacement of all coal in electricity generation with renewables (e.g. hydro and solar power) and nuclear by 2050. In the 2DS scenario, a majority of the reductions in GHG emissions is gained through switching away from coal use and through the use of CCS in the power generation sector, but changes affect all other sectors as well. Both of these scenarios would require large investments into new energy technologies and infrastructure, including updates to the electricity grid to integrate increased share of renewables. The 2007 LTMS report suggests that these changes are feasible, but would require immediate investment in research and development into the required technologies (ERC, 2007).

Both of the climate change mitigation scenarios have a number of air pollution emissions co-benefits. However, past studies in South Africa and elsewhere have provided evidence that increasing electricity prices due to large investments by the electric utility will cause poorer communities to use less electricity for cooking and heating (Madubansi & Shackleton, 2007; OECD, 2012). Instead, these households will turn to cheaper solid fuels (coal, charcoal, and wood), which are the main contributors to particulate exposure. The *RENewables Tradeoff* (RENT) scenario is modeled using the same replacement of coal with renewables and nuclear as in the renewable scenario, but with a tradeoff of a 10% reduction in electricity use in the domestic sector and a corresponding increase in domestic solid fuel use. Previous studies have not predicted the extent to which increasing electricity prices will affect electricity use in the domestic sector, so this study provides a stylized example.

Activity inputs for each of the activity scenarios (supplemental, S2) show the absolute differences in energy consumption by fuels, and how consumption is projected to change over time. The scenarios that involve the adoption of significantly more efficient technologies (notably CCT, 2DS, REN and RENT) require less energy inputs than the others. For example, the CCT scenario, which involves a switch to highly efficient IGCC power plants, requires less coal input to produce the same energy. Activity inputs for the 2DS scenario show the shift from coal toward biomass, renewables, and nuclear.

8.5 Results

8.5.1 Business as usual scenario

Emissions of NO_X and CO_2 are projected to grow in 2030 and 2050 under the BAU scenario (Figure 8-3). $PM_{2.5}$ emissions will remain relatively constant until 2050, while SO_2 emissions will decrease by one third by 2030 and remain relatively stable until 2050. Projected reductions in $PM_{2.5}$ and SO_2 emissions in the power generation sector are the cause of tightening regulations on emissions, and the realization of these reductions will depend on the ability of the power and industrial sectors to install the controls necessary to meet these standards.

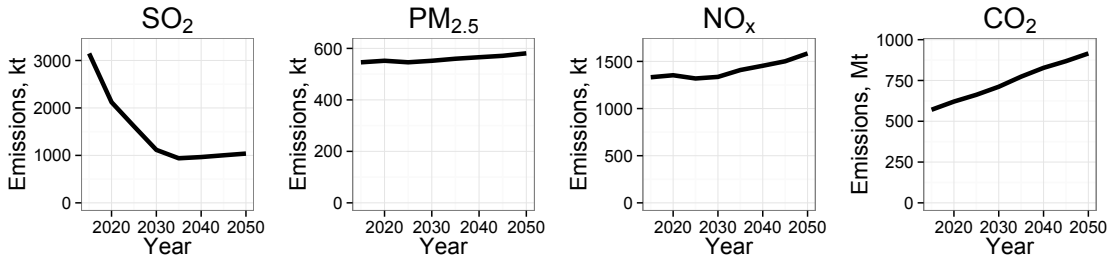


Figure 8-3. Estimated emissions under the BAU scenario.

8.5.2 Control scenarios

The two control scenarios show the potential for removing pollutants without changing the activity mixture in the BAU scenario (Figure 8-4). For example, SO_2 emissions are projected to decrease by over 50% in the BAU scenario in 2030 compared to 2015. By not investing in further SO_2 controls (the NFC scenario), South Africa would see its SO_2 emissions increase due to projected reliance on a fossil fuel driven energy system. Further, an investment in the best available technology to control SO_2 (the MFC scenario) would decrease SO_2 emissions by another 50% in 2030 compared to BAU.

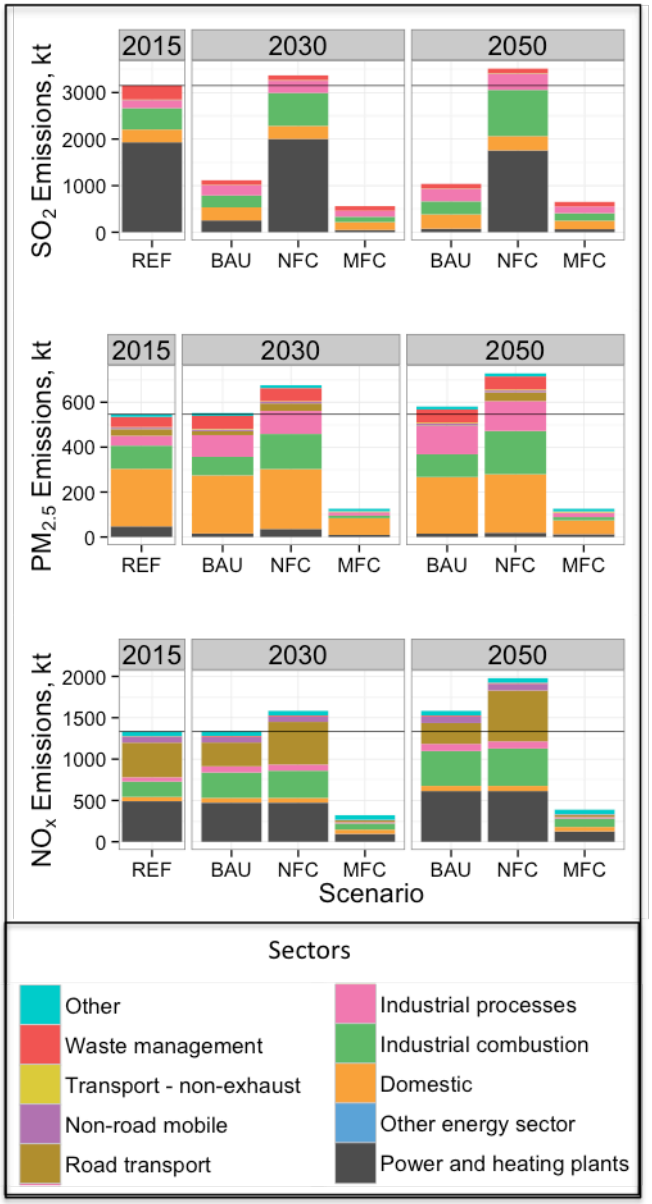


Figure 8-4. Emissions in the BAU and control (NFC and MFC) scenarios. REF refers to the estimated emissions for 2015 used as a baseline in all simulations. The horizontal line in each plot corresponds to estimated total 2015 emissions.

Results are considerably different for $PM_{2.5}$ and NO_x. The BAU scenario for these pollutants yields emissions that are much closer to the NFC scenario emissions than the MFC scenario, which would yield a reduction in $PM_{2.5}$ emissions by 77% and NO_x emissions by 75% in 2030 compared to BAU. Therefore, current legislation requires emissions reductions that are much less than potential reductions if investments were made in improved removal technologies. $PM_{2.5}$ emissions from the electricity generation sector

are expected to decrease in the all scenarios largely because of the retirement of older coal plants.

South Africa has a large potential to reduce NO_x emissions in the road transportation sector. Estimated emissions in the BAU case change very little between 2030 and 2050, while implementing maximum controls (which correspond to a shift to EURO 6 engines and fuels) would reduce NO_x emissions by 89%. $PM_{2.5}$ emissions from road transportation are also projected to decrease in the BAU scenario by 2050, although controlling emissions to their fullest extent in the MFC scenario would yield emissions that are 32% of BAU emissions.

8.5.3 Policies that affect domestic air pollution emissions

Domestic activities are projected to contribute the largest share of $PM_{2.5}$ emissions by 2030 in the BAU scenario (Figure 8-4). These emissions are especially important from a public health perspective, as exposure to indoor ambient air pollution is very high for those who cook and heat their homes with solid fuels. The DOM scenario, in which domestic use of solid fuels is replaced with LPG, reduces $PM_{2.5}$ emissions in this sector by 75% by 2030 compared to 2015 (Figure 8-5), just less than the MFC control scenario. NO_X emissions are reduced by a third in the domestic sector, but this reduction is small in comparison to the total NO_x emissions.

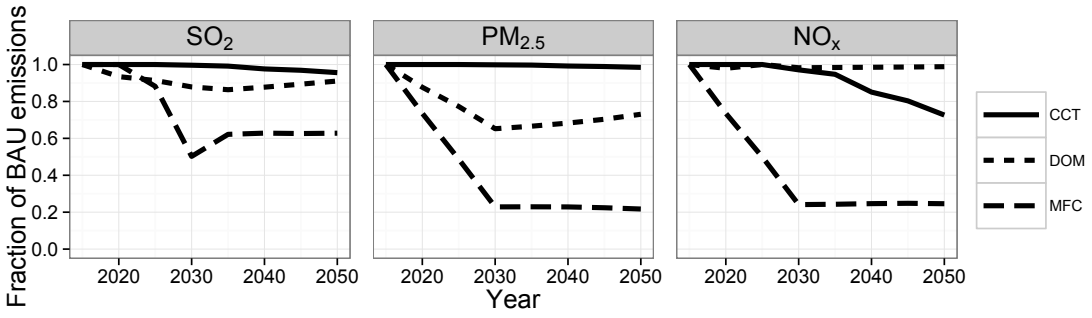


Figure 8-5. Emissions in the air pollution reduction scenarios (CCT, DOM, and MFC) scenarios relative to BAU.

Domestic emissions in the REN scenario, which affects emissions only in the power generation sector, do not deviate from the BAU scenario (Figure 8-6). If a switch to renewables causes the 10% decrease in demand for electricity as modeled in the RENT scenario, domestic $PM_{2.5}$ emissions will increase by 10% in 2030 and 16% in 2050.

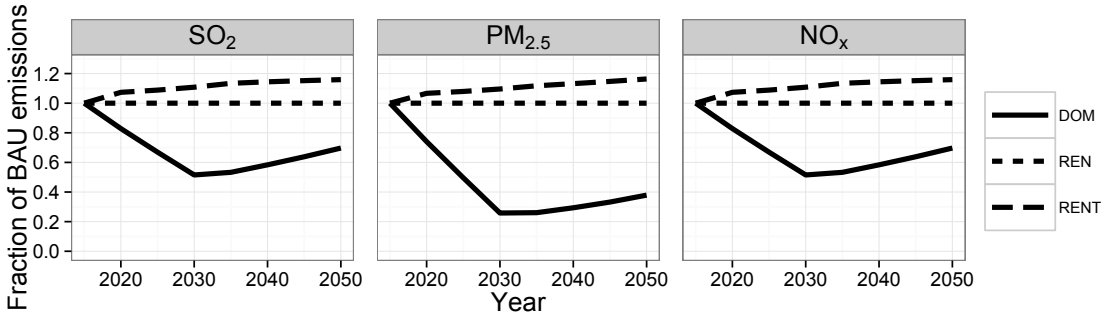


Figure 8-6. Domestic emissions relative to BAU for the DOM, REN, and RENT scenarios.

It is beneficial to compare emissions reductions possible through control scenarios and scenarios that target the dirtiest energy production activities (Figure 8-5). When compared to the MFC option, activity options are less effective and slower to implement. For example, controls have the potential to reduce SO_2 emissions by nearly 1,500 kt, $PM_{2.5}$ by more than 500 kt, and NO_x by more than 1,000 kt in 2030. The two other scenarios aimed at decreasing local air pollutant emissions (CCT and DOM) are not able to remove as much as the MFC scenario, although it is important to note that a combination of these scenarios would produce additive reductions since the two target different activity sectors.

8.5.4 Climate change mitigation scenarios

Three scenarios (CCT, 2DS, and REN) have a substantial effect on both GHGs and air pollution emissions (Figure 8-7). Of the three scenarios, the 2DS scenario reduces emissions of all four pollutants most rapidly. The CCT and REN scenarios do not reduce emissions before 2025 because both are designed to allow existing power plants to stay in operation until their projected retirement date. The CCT scenario reduces NO_x and CO_2 emissions to a stable level of near 50% of BAU by 2040, a value that the REN scenario does not reach until 2050. Further, the 2DS scenario sees the largest benefits for SO_2 and $PM_{2.5}$.

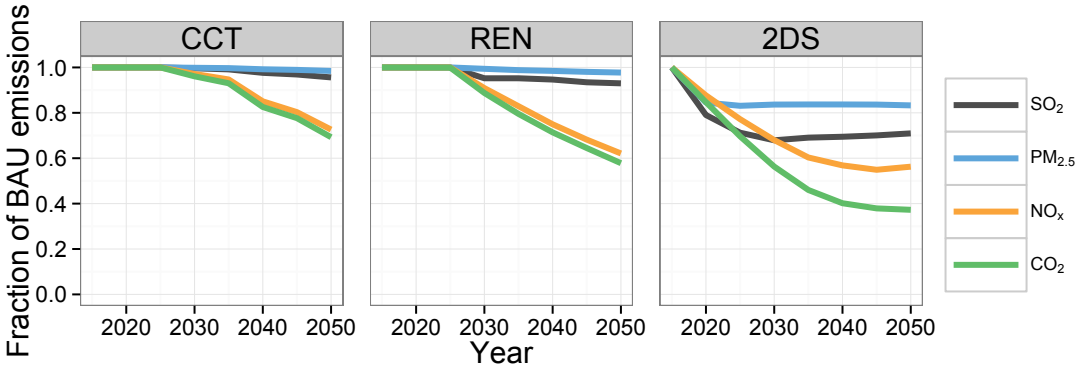


Figure 8-7. Total emissions relative to BAU for the CCT, 2DS, and REN scenarios.

8.5.5 Control costs

GAINS estimates the costs of controls installed for each scenario using a 4% interest rate (Figure 8-8). These costs can be compared to projected GDP to help understand the price in the context of South Africa's total economy (Table 8-2). As expected, the MFC scenario would require the greatest investments in control technologies, with costs totaling 1.8% of GDP in 2030. Costs are avoided in the climate scenarios (2DS and REN) by switching to energy producing activities that do not require air pollution controls (e.g. solar and wind energy).

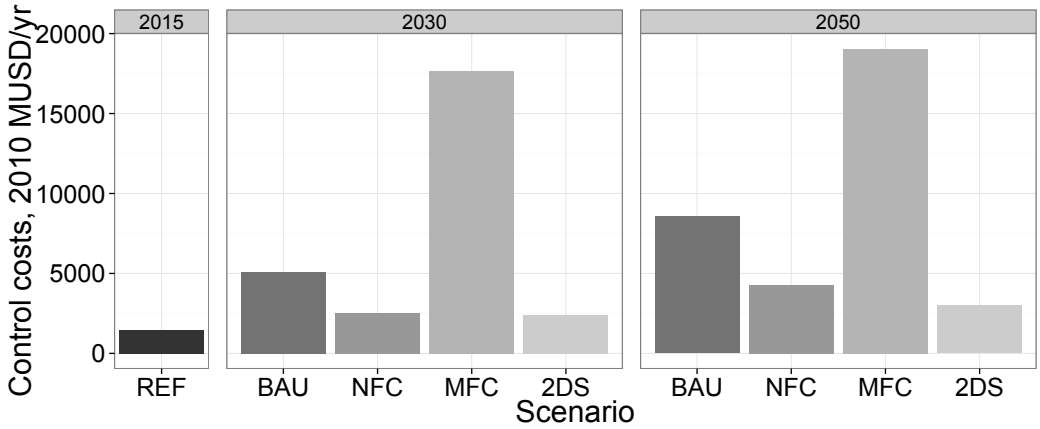


Figure 8-8. Costs of emissions controls for selected scenarios in 2010 USD. REF represents the estimated costs for 2015.

Table 8-2. Control costs for selected scenarios as a percent of GDP.

Scenario	2030	2050
BAU	0.5	0.5
NFC	0.2	0.2
MFC	1.8	1.1
2DS	0.3	0.2
REN	0.4	0.2

8.6 Discussion

8.6.1 Evaluation of GAINS output compared with previous studies

A comprehensive bottom-up inventory of South Africa's local air pollution emissions does not exist, although there are plans to develop one (see *www.saaqis.org.za/Emissions3.aspx*, accessed 11 November, 2015). Four sources provide a point of comparison for the emissions estimated by GAINS (Table 8-3). The first, the Emission Database for Global Atmospheric Research (EDGAR), is a joint effort under the European Commission to assemble a database of global air pollution and GHG emissions on a grid (EC, 2011). The second is a report compiled for the South African Department of Environmental Affairs (DEA), and presents GHG emissions estimates in South Africa from 2000 to 2010 (Witi et al., 2013). Seymore et al. (2014) compared multiple estimates, and was therefore able to provide an estimate of uncertainty. The earliest year available from this report 2007, is presented alongside the GAINS 2005 value. Van der Hoeven (2011) primarily based the estimates on fuel combustion using the 1996 IPCC guidelines. 2005 is used to compare emissions because this year is available from a majority of the sources.

Table 8-3. Comparison of GAINS-estimated emissions to previous estimates in 2005.

Source	SO_2 (kt)	NO_x (kt)	CO_2 (Mt)
GAINS–BAU	2800	1300	420
EDGAR (EC, 2011)	2200	1400	360
Van der Hoeven, 2011	–	–	330
GHG-SA (Witi et al., 2013)	–	–	520
Seymore et al., 2014	–	–	465 ± 38[a]

[a]Estimate is for 2007

GAINS estimates SO_2 and NO_x emissions to within 30% and 9% of the EDGAR values, respectively (Table 8-3). Major categories, such as electricity production plants, domestic, and transportation emissions match well with the EDGAR estimates. Industrial processes and other sources in GAINS make up the bulk of the difference in SO_2 emissions. National estimates of $PM_{2.5}$ emissions were unavailable to compare to GAINS estimates.

The GAINS estimate of CO_2 emissions in 2005 falls squarely in the middle of the four studies used for comparison. A reason for the higher estimate from the DEA GHG inventory is that the authors estimate nearly 20% greater emissions for power generation and slightly higher for each of the other sectors in the energy sector (transportation, domestic energy production, etc.). In general, GAINS estimates higher emissions than the methods that rely mostly on fuel consumption data (EDGAR & Van der Hoeven), and somewhat lower emissions than the other two methods.

8.6.2 Air pollution reduction policies

Both air pollution control scenarios and energy activity scenarios provide pathways for reducing air pollution emissions in South Africa (Figure 8-3 through Figure 8-6). Results from the MFC scenario show that the country has the potential to reduce emissions of SO_2, NO_X, and $PM_{2.5}$ from all sectors by investing in control options. Controlling emissions from existing sources provides an opportunity for industries to extend the lifetime of current facilities and meet emissions standards. There are costs associated with installing controls, but these costs delay the costs of implementing renewables and other cleaner technologies. However, control options must be considered along with other options (e.g. renewable sources of energy) in the face of limited resources.

Replacing solid fuels with LPG in the domestic sector has as large an effect on $PM_{2.5}$ emissions reductions as costly controls installed in the industrial and power sectors. Further, emissions reductions in the domestic sector have a large effect on exposure. These policies (i.e. replacing all solid fuel use with clean energy forms) would not come without costs and investments in infrastructure, which are not estimated within the GAINS model.

The CCT scenario option has a number of benefits to South Africa. For instance, emissions of NO_x from the electricity-generating sector decrease significantly (Figure 8-7). Further, CO_2 emissions are also reduced compared to the baseline due to gains in conversion efficiency. If carbon capture and storage technologies were developed to a

financially viable level, IGCC plants with CCT installed might provide an effective emissions-reduction option for multiple pollutants. Although this technology is costly, such plants may be a necessary option that allows South Africa to encourage economic growth and reduce emissions while continuing to exploit its abundant coal resources.

8.6.3 Climate change mitigation policies

Climate change policies are developed to contribute to the share of global emissions that must be cut if the world is to remain below a certain level of temperature rise caused by increasing concentrations of GHGs in the atmosphere. Many of the policies, however, also carry the benefits associated with burning less fossil fuel, including reduced local air pollution emissions and associated negative health effects. Costs that are normally associated with implementing technologies (such as emissions controls) to reduce air pollution can instead be invested in renewables or other energy sources with the same result in decreases in emissions.

The 2DS scenario is the most ambitious of the climate scenarios in terms of estimated CO_2 emissions. A point of interest is that the scenario does not completely eliminate emissions from coal-fired power plants, instead targeting all sectors. This is in contrast with the REN scenario, which eliminates coal use in electricity production. The 2DS is closest to the scenario presented in the LTMS study, which the authors reported will be difficult to achieve even if South Africa begins to adopt aggressive climate mitigation policies immediately. The REN scenario, therefore, shows an alternative approach that targets only one sector. CO_2 emissions are not projected to decrease as much in this scenario (42%) as the 2DS scenario (54%) in 2050 compared to BAU, but it provides an option that allows other sectors to function as they would compared to BAU.

8.6.4 Policy co-benefits and contradictions

An advantage of the GAINS modeling platform is that emissions of many air pollutants can be assessed simultaneously for single scenarios. Emissions of GHGs can be estimated for scenarios that are designed to control local air pollutant emissions. The CCT scenario provides one example of this. Even though the scenario is designed to reduce emissions of local air pollutants usually inherent in coal burning (e.g. SO_2, $PM_{2.5}$, and NO_x), results show that CO_2 emissions are also reduced in this scenario compared to BAU (Figure 8-7). In fact, reductions in relative NO_x and CO_2 emissions reductions far exceed

reductions in SO_2 and $PM_{2.5}$ emissions for the CCT scenario. Comparative reductions in SO_2 and $PM_{2.5}$ appear lower because current regulations in South Africa require controls in the BAU scenario. Climate change policies (2DS and REN) also yield reductions in local air pollution emissions. The largest reductions for all four pollutants are achieved in the 2DS scenario.

A danger of adopting aggressive climate change policies, however, is the increasing electricity prices that accompany large investments in new technologies (Ürge-Vorsatz et al., 2012). Newly electrified poorer communities tend to use electricity to supplement their use of solid fuels, usually for lighting and entertainment (e.g. televisions or charging cell phones) (Madubansi & Shackleton, 2006). Deloitte (2011) reported that it is difficult to estimate the effect of electricity pricing on demand, and historically the demand for electricity has been more dependent on income (or GDP) than on price. However, they note that if prices increase beyond a threshold this may change. The RENT scenario attempts to capture this effect by reducing the demand for electricity in the domestic sector by 10% and increasing the energy input of solid fuels by a corresponding amount. Results from this scenario show that even though reductions in emissions of $PM_{2.5}$ are realized by adopting the aggressive RENT scenario, emissions from the domestic sector increase by 16% in 2050 (Figure 8-6). Multiple studies have shown that in South Africa, indoor air pollution is the most important contributor to exposure (Scorgie et al., 2003), meaning this increase would translate into a disproportional increase in negative health outcomes in densely populated poor communities.

Although climate policies have a positive effect on reductions of GHGs in power generation, they must be accompanied with other policies (e.g. subsidies to encourage use of electricity or LPG for cooking) to ensure the strategies do not cause negative effects on emissions from other sectors. These co-benefits and tradeoffs need to be investigated further to ensure the maximum returns are achieved when designing policies to address air pollution and climate change.

A further benefit is the avoidance of costs of controlling air pollutants (Figure 8-8). The REN and 2DS scenarios are projected to have lower control costs than the BAU scenario in both 2030 and 2050. In 2050, both scenarios will have total costs of less than the NFC scenario, meaning that changing the activity mix has just as great an effect on

control costs as simply maintaining controls at the current levels. These avoided costs serve as a source of savings in otherwise costly climate mitigation scenarios.

8.6.5 Policy frameworks for achieving goals in the scenarios

A companion paper (Klausbruckner et al., 2016) assesses the current energy and associated environmental policies in South Africa in detail. The authors identify contradictions in the current framework, and discuss opportunities for an exploitation of co-benefits of an integrated approach to managing energy production and use, air pollution, and climate change mitigation policies. They discuss multiple policy tools that could be used to achieve the scenarios detailed above. For example, carbon taxes would encourage a shift towards more of the technologies that are necessary to achieve targets simulated in the CCT, 2DS, or REN scenarios. Such a policy could be further enhanced by offering transfer payments to lower-income households to shield them from price increases that would accompany a large investment into new energy production technologies.

South Africa has enacted regulations to enforce emissions limits from industrial facilities and power plants. To encourage facilities to further reduce emissions through, for example, control retrofits, the country has a number of policy options and examples of their implementation in other regions. For example, SO_2 and NO_X cap-and-trade marketplaces have generally been shown to be effective in the United States (Ellerman, 2003; Morgenstern et al., 2012).

In the transportation sector, one limitation to shifting towards advanced control technologies (e.g., the Euro 6 standard) is the continued use of high sulfur fuels, which are incompatible with Euro-6 compliant engines. Further, the persistence of older automobiles limits the penetration of new technologies.

Klausbruckner et al. conclude that the most efficient approach to developing policy must include an integrated assessment of its effects on multiple sectors of the economy (as a whole and on historically disadvantaged communities), public health, and air pollution and GHG emissions. The current work applies GAINS in such an integrated assessment, though the results of the preliminary model are limited in scope and leave room for a more detailed assessment of each scenario.

8.7 Conclusion and Policy Implications

South Africa's energy intensive economy emits large amounts of air pollution that contributes both to negative health and environmental effects locally and climate change on a global level. The GAINS model has been used to assess the impacts of a number of policy options that could be pursued by the country to both reduce air pollution and mitigate GHGs.

Scenarios that target GHG emissions also reduce SO_2, $PM_{2.5}$, and NO_X emissions by switching away from an economy based on fossil fuels, which would translate into lower incidences of negative health impacts caused by air pollution. Further, costs are saved on controlling emissions from conventional technologies (e.g. coal combustion) used to produce electricity.

Care must be taken when implementing policies to encourage controls and changes in activities, since rising prices may encourage poorer communities to switch away from electricity use in their households to more polluting solid fuels. More work is necessary to study how this tradeoff will change with increasing prices and changing income levels. Further development of the GAINS integrated modeling framework will require implementation of parallel dispersion modeling, exposure, and health effects models, and valuation of the public health influences of various scenarios. These adjunct models will increase the value of this type of integrated systems analysis approach as an evidence-based policy support mechanism in reaching decisions in increasingly complex systems of energy, economics, environmental and public health.

8.8 Acknowledgements

This work has been completed as part of the International Institute for Applied Systems Analysis (IIAASA)'s 2014-15 Southern African Young Scientists Summer Program organized by South Africa's National Research Fund and Department of Science and Technology and the University of the Free State, Bloemfontein, South Africa.

8.9 References

Amann, M., Bertok, I., Borken-Kleefeld, J., Cofala, J., Heyes, C., Höglund-Isaksson, L., … Winiwarter, W. (2011). Cost-effective control of air quality and greenhouse gases in Europe: Modeling and policy applications. *Environmental Modelling & Software*, *26*(12), 1489–1501. https://doi.org/10.1016/j.envsoft.2011.07.012

Amann, M., Kejun, J., Jiming, H., Wang, S., Wei, W., Jia, X., … Winiwarter, W. (2008). *Scenarios for cost-effective control of air pollution and greenhouse gasses in China*. Retrieved from http://gains.iiasa.ac.at/gains/download/GAINS-Asia-China-v8_MA.pdf

Bentley West Management Consultants (Pty) ltd. (2004). *Fund for reserach into industrial development growth and equity (FRIDGE)*.

Department of Local Government and Housing. (2010). *Gauteng Integrated Energy Strategy*. Marshalltown, Gauteng, South Africa.

Edkins, M., Marquard, A., & Winkler, H. (2010). *South Africa's renewable energy policy roadmaps*.

Ellerman, A. D. (2003). *Ex Post Evaluation of Tradeable Permits :* (MIT-CEEPR (Series) ; 03-003WP.). Retrieved from http://hdl.handle.net/1721.1/44996

European Commission. (2011). Emission Database for Global Atmospheric Research. https://doi.org/10.2904/EDGARv4.2

IIASA. (2012). GAINS Global. Laxenburg, Austria. Retrieved from http://gains.iiasa.ac.at/models/index.html

International Energy Agency. (2012). *Energy Technology Perspectives 2012*.

IPCC. (2001). *Climate Change 2001: Synthesis Report*. Retrieved from http://www.grida.no/publications/other/ipcc_tar/?src=/climate/ipcc_tar/

Jongikhaya Witi, Stevens, L., Manzini, L., & Musee, N. (2013). *GHG Inventory for South Africa*.

Kiratu, S. (2010). *South Africa's Energy Security in the Context of Climate Change Mitigation*.

Klausbruckner, C., Annegarn, H. J., Henneman, L. R., & Rafaj, P. (2016). A policy review of synergies and trade-offs in South African climate change mitigation and air pollution control strategies. *Environmental Science & Policy*, *57*, 70–78. https://doi.org/10.1016/j.envsci.2015.12.001

Lesiv, M., Bun, A., & Jonas, M. (2014). Analysis of change in relative uncertainty in GHG emissions from stationary sources for the EU 15. *Climatic Change*, *124*(3), 505–518. https://doi.org/10.1007/s10584-014-1075-6

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., … Ezzati, M. (2013). A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*, *380*(9859), 2224–60. https://doi.org/10.1016/S0140-6736(12)61766-8

Long Term Mitigation Scenarios Technical Summary. (2007).

Madubansi, M., & Shackleton, C. M. (2006). Changing energy profiles and consumption patterns following electrification in five rural villages, South Africa. *Energy Policy*, *34*(18), 4081–4092. https://doi.org/10.1016/j.enpol.2005.10.011

Madubansi, M., & Shackleton, C. M. (2007). Changes in fuelwood use and selection following electrification in the Bushbuckridge lowveld, South Africa. *Journal of Environmental Management*, *83*(4), 416–26. https://doi.org/10.1016/j.jenvman.2006.03.014

Merven, B., Moyo, A., Stone, A., Dane, A., & Winkler, H. (2014). *Socio - economic implications of mitigation in the power sector including carbon taxes in South Africa*.

Morgenstern, R. D., Harrington, W., Shih, J.-S., & Bell, M. L. (2012). Accountability Analysis of Title IV Phase 2 of the 1990 Clean Air Act Amendments, (168).

Myllyvirta, L. (2014). *Health impacts and social costs of Eskom's proposed non-compliance with South Africa's air emission standards*. Exeter, UK.

OECD. (2012). *OECD Environmental Outlook to 2050: The Consequences of Inaction*. Paris. Retrieved from http://www.keepeek.com/Digital-Asset-Management/oecd/environment/oecd-environmental-outlook-to-2050_9789264122246-en#page297

Pauw, C., Friedl, A., Holm, D., John, J., Kornelius, G., Oosthuizen, R., & Njekerk, A. van. (2008). *Air pollution in dense, low-income settlements in South Africa*.

Purohit, P., Munir, T., & Rafaj, P. (2013). Scenario analysis of strategies to control air pollution in Pakistan. *Journal of Integrative Environmental Sciences*, *10*(2), 77–91. https://doi.org/10.1080/1943815X.2013.782877

Rafaj, P., Schöpp, W., Russ, P., Heyes, C., & Amann, M. (2013). Co-benefits of post-2012 global climate mitigation policies. *Mitigation and Adaptation Strategies for Global Change*, *18*(6), 801–824. https://doi.org/10.1007/s11027-012-9390-6

SA Department of Energy. (2013). *Integrated Resource Plan for Electricity 2010-2030*.

SAPIA. (2014). Future Fuel Specifications. Retrieved March 8, 2015, from http://www.sapia.co.za/key-issues/environment.html

Scenario Building Team. (2007). *Lont Term Mitigation Scenarios Strategic Options for South Africa*. *Long Term Mitigation Scenarios: Scenario Document, Deaprtment of Environment and Tourism*.

Scorgie, Y., Kneen, M., Annegarn, H. J., & Burger, L. (2003). Air Pollution in the Vaal Triangle - Quantifying Source Contributions and Identifying Cost-Effective Solutions. *Clean Air Journal*, *13*(2), 5–18.

South African Government. (2011). *National Climate Change Response White Paper*. Pretoria, SA. Retrieved from http://www.climateresponse.co.za/

Tait, L., & Winkler, H. (2012). *Estimating greenhouse gas emissions associated with achieving universal access to electricity in South Africa*.

U.S. Energy Information Administration. (2013). South Africa. Retrieved March 8, 2015, from http://www.eia.gov/countries/country-data.cfm?fips=SF#cde

Van der Hoeven, M. (2011). CO2 Emissions from Fuel Combustion Highlights. *International Energy Agency*, 1–134. https://doi.org/10.1787/co2_fuel-2011-en

Wagner, F., Schöpp, W., & Amann, M. (2013). Dealing with fixed emissions ceilings in an uncertain future: Offsetting under environmental integrity. *Journal of Environmental Management*, *129*, 25–32. https://doi.org/10.1016/j.jenvman.2013.05.054

Winiwarter, W. (2005). *The GAINS Model for Greenhouse Gases – Nitrous Oxide*.

Winkler, H. (2007). Energy policies for sustainable development in South Africa. *Energy for Sustainable Development*, *XI*(1).

World Nuclear Association. (2014). Nuclear Power in South Africa. Retrieved from http://www.world-nuclear.org/info/country-profiles/countries-o-s/south-africa/

CHAPTER 9. CONCLUSIONS AND FUTURE RESEARCH

This thesis has presented detailed assessments of impacts of regulations on emissions and air quality with an emphasis on policy implications. The bulk of the work focuses on past air quality policies in the United States, and results of the accountability assessment, overall, are positive. Even when accounting for numerous extraneous factors that interjected uncertainty into relationships along the accountability chain, regulations on utilities and mobile sources in the United States successfully reduced emissions and improved air quality. Much of the research presented here is published or in preparation for submission, including in a report to the Health Effects Institute (Russell et al., 2017) and in multiple refereed journal articles (Henneman, Holmes, et al. 2015; Henneman, Rafaj, et al. 2015; Henneman et al. 2016; Henneman et al. 2017; Henneman et al., n.d.). This chapter summarizes the conclusions from the previous chapters and offers suggestions for future research.

9.1 Conclusions

The following conclusions are divided into three areas: *accountability, air quality in the United States, and South Africa energy, air quality, and climate*. Each of these reflect an important area of focus of the preceding research, and are bound by an interest in quantifying impacts of policies.

9.1.1 Accountability

Air quality controls are expensive, but have been shown to save numerous lives and improve public and environmental health. These outcomes can help justify the costs, and highlight the need for a structured approach to assess the efficacy of regulatory progress, i.e., accountability research. Intuitively, control programs should reduce emissions, improve air quality, and lead to fewer negative health and environmental effects; however, since regulations are implemented concurrent with innumerable changes across the Accountability Chain (Health Effects Institute 2003), quantifying with certainty causal relationships along the chain remains arduous. Mixed results of previous studies underline the importance of carefully defining the intervention of interest and considering as many confounding factors as possible while following the regulation's signal along the Chain.

The accountability portion of the study benefited from multiple factors. Major emissions reductions associated with the controls had large impacts on emissions and air quality; more modest regulations would have made resolving the differences between actual and counterfactual more difficult. Multiple detailed analyses throughout the work benefited from largescale data availability (e.g., high quality air quality and emissions monitoring networks) and models tailored for air quality in the United States (e.g., meteorological modeling, mobile source emissions modeling, and chemical transport modeling). Regional stakeholders and other experienced professionals provided invaluable information on the history of air quality regulations and practical matters relating to their implementation. Their contributions supplemented vast amounts of publicly available information on the regulatory frameworks in the United States.

Although these tools and data are not available for all accountability studies, the framework applied here of careful planning and accounting for confounding factors is unique, and will prove valuable for future policy assessments.

9.1.2 Air quality in the United States

Work showed that regulatory impacts on ambient concentrations are source-, species-, and seasonally-dependent. Mobile source controls have had relatively greater impact on O_3 concentrations than large stationary sources. Large stationary sources have had greater impact on $PM_{2.5}$ than mobile sources, primarily due to their contribution to sulfate from coal plants. Meteorological variability was shown to be an important contributor to daily concentration variability, especially for O_3, but was less important for long-term trends of all pollutants. Median concentration changes between 2001 and 2011 in the United States were attributable almost exclusively to emissions changes (although for ozone, median changes are much smaller than changes at the extremes).

CMAQ, an air quality model used in regulatory and scientific applications, was shown to imperfectly reproduce O_3 and $PM_{2.5}$ concentrations and changes between 2001 and 2011. Sensitivities to emissions and meteorology for these species agreed reasonably with empirical estimates in the southeastern United States. CMAQ concentration and sensitivity simulations of other gaseous and particulate species had higher bias and error than O_3 and total $PM_{2.5}$. CMAQ-estimated ozone production efficiency (OPE), a metric related to the potential impacts of emissions controls, is lower than empirical-modeled OPE

in urban areas, but not in rural areas. Ozone production efficiency has increased over recent decades with emissions reductions, but NO_X emissions controls remain a viable option for lowering the highest O_3 concentrations in the Southeast.

9.1.3 South Africa energy, air pollution, and climate

The South Africa work (Chapter 8) produced multiple outcomes. First, the application of GAINS, an integrated assessment model, allowed for the assessment of multiple policies simultaneously under the same modelling framework. With consistent model assumptions and inputs, South African emissions and controls costs could be compared directly. Results showed numerous viable policy options that benefit energy, air pollution, and climate. GAINS model results showed co-benefits and tradeoffs of various policies, such as reductions in emissions control costs that accompany a switch to renewables. The approach provides supporting evidence for policies that exploit co-benefits and avoid contradictions by assessing multiple aspects of the energy sector within an integrated framework.

9.2 Future research

At its core, all research in the air quality field is driven by two major motivators: public and ecological health. Looking forward, there is no reason to believe this will change. This Future Research discussion focuses on needs in three areas—*accountability, air quality modeling,* and *future policy projections*—that are motivated by the two driving factors and will build off the research foundation laid by this thesis.

9.2.1 Accountability

Over the fifteen years since the Health Effects Institute's report on the need for accountability studies (Health Effects Institute 2003), the field has expanded dramatically. The expansion has coincided with improvements in air quality in developed countries and methods with which to quantify these improvements. Even with these tools, accountability studies have met limited success at causally attributing improvements in air quality and health to regulatory actions.

Future successful accountability studies, besides taking advantage of increased availability of observations and modelling capabilities, should leverage the expertise of researchers in multiple fields. Multidisciplinary collaborations provide an opportunity in accountability research to assess the complete impact of regulations (and more confidently

project impacts forward). For example, cost, a major motivator of accountability assessments, was not considered in this thesis. Economists have developed advanced methods for evaluating costs and benefits associated with various public health and environmental outcomes (e.g., U.S. EPA 2011) and would therefore make valuable collaborators for future accountability assessments.

An immediate opportunity for an accountability assessment is to adapt the empirical methods presented in Chapters 3, 4, and 5 for implementation across the entire United States. While this work (Chapter 6) and others estimate emissions/meteorological impacts on air pollution concentrations, the creation of counterfactual concentration fields would aid regulators in comparing effectiveness of certain control strategies (as it did in the studies of Atlanta air quality (Russell et al., 2017)) and would support epidemiologists in assessing health impacts of specific controls.

The expanded study would require additional data sources, model runs, and updated statistical and data handling methods. Ambient air quality and meteorological observations will come from the EPA's Air Quality System (AQS—http://www.epa.gov/ttn/airs/airsaqs/), which contains data for the over 10,000 monitors active over the period 1990-2017. EGU emissions data will again come from the EPA's Air Markets Program Data (AMPD—http://ampd.epa.gov/ampd/) archive, which houses daily emissions at the facility level. MOVES, EPA's MOtor Vehicle Emissions Simulator (https://www.epa.gov/moves), will again model daily county-level mobile source (i.e., automobile) emissions. These major emissions sources will be supplemented as necessary with other major sources in some locations with interpolated emissions from the National Emissions Inventory (https://www.epa.gov/air-emissions-inventories/national-emissions-inventory-nei), which is released every three years.

The major innovation from this work will be a national statistical framework linking emissions sources and observed ambient air pollution concentrations that accounts for regional differences in meteorology, transport, and chemistry. The proposed model would offer an alternative to previous statistical and chemical transport model approaches characterizing air quality on a large scale that is more applicable to accountability research, but would also benefit from lessons learned and comparisons with these previous modeling approaches.

The proposed study will produce two important outcomes: daily source-specific air quality impact estimates on monitors throughout the United States and counterfactual concentrations estimates. Estimates will benefit from being based on observations, instead of being exclusively modeled (for example, using a chemical transport model), and will include uncertainty estimates. Spatial and temporal differences in daily air quality responses to emissions changes and their uncertainties will inform future policy development, and a commitment to involving stakeholders in model development will enhance the quality and communication of the results.

9.2.2 Air quality modelling

All models—which codify our knowledge of relationships in systems—are wrong (Oreskes et al. 1994). We cannot develop complete models because we are unable to observe all aspects of a system. Far from diminishing models' utility, this fact means simply that researchers must evaluate each model's suitability to specific problems before they draw conclusions from their results.

With the evolution and growing abundance of new observation platforms such as widespread monitoring networks, high-cost—but very powerful—research instruments, low-cost sensors, satellites, and emissions monitoring equipment, the atmospheric chemistry community is dramatically broadening its observation capacities. While each technology has drawbacks (some of which are not fully characterized), the suite of options greatly increases the amount and types of data available for inferring relationships in the physical world. These observations will help in both model development (e.g., further improvements in organic and sulphate chemistry in CMAQ (Chapter 6) and ozone responses to emissions changes in cities (Chapter 7)), and in application-specific model evaluations.

It is straightforward to list a variety of policy-relevant air quality model applications (e.g., modelling absolute concentrations, concentrations changes over time, impacts of future controls, performance at low concentrations, and model sensitivities to emissions sources, variability at different time/distance scales (hourly, daily, annual, decadal; 10s of meters to 1000s of km), and specific meteorological conditions (e.g., low wind speed). Models for each of these applications have been developed, but come with their own limitations. Each model implementation implies its own evaluation; for example, CMAQ

performs better for MDA8h ozone (the regulated value) than hourly ozone, and reproduces the change in MDA8h ozone between 2001 and 2011. In this case, modeling for the regulatory application is more appropriate for the model than for hourly ozone.

Large amounts of new data require the tailoring of data handling and statistical techniques to handle and analyse vast datasets (work related to this thesis, for example, filled more than 64 TB of disk space). Potential applications of these data include combining observation and modelled data using various techniques, including detailed model evaluations (including the work in Chapter 6), data fusion approaches (Friberg et al. 2016; Hu et al. 2014), and land use regression (de Hoogh et al. 2014; Kloog et al. 2012). Besides serving as a target for deterministic models, observations drive empirical modeling, which in turn aids deterministic model formation.

Future work in the field should leverage lessons learned and interdisciplinary collaborations to further improve the combinations of modeled and observed data. For example, dense sensor networks could be used to increase the resolution of chemical transport modeled air quality fields in areas with dense populations. In areas with dense monitoring networks, approaches such as this could be adapted into a dynamic air quality management system to help city planners warn the public of poor air quality days. Projects such as this are important for understanding natural and anthropogenic impacts on pollutants with high spatial variability, such as $PM_{2.5}$.

9.2.3 Policy projections

Improvements in air quality accountability and modelling culminate in successful projections of proposed air quality regulations (National Academies of Sciences Engineering and Medicine 2016). The most important difference between accountability studies and projections is a logistical one—projections require the development of both the baseline (often termed ‘business as usual’) and with-regulation scenarios, whereas in accountability studies the with-regulation scenario is known (Cropper et al. 2017).

Projections in developed and developing countries will occur in two different pollution regimes, and will therefore rely on separate experience bases. In general, policy projections for developed countries will be made at lower ambient concentrations. In these conditions, background and imported concentrations grow in importance along with non-anthropogenic air pollution sources such as wildfires and dust (although these are

influenced by anthropogenic activities). Projections in developing countries will rely on previous experience of developed countries in high-pollution regime, though their individual policies will need to be designed based on local conditions—e.g., meteorology, energy production profile, and technological feasibility. Policy projections will all require accounting for added uncertainty of a changing climate and its various feedback loops and co-effects (e.g., Fiore et al. 2015).

Integrated assessments, such as the GAINS work (Chapter 8), provide consistent frameworks in which to assess policy implications across a range of disciplines, and require interdisciplinary collaboration. Integrated assessment models combine individual models, for instance, relating electricity use to emissions, emissions to air quality, and air quality to health. Most applications to date use relatively simplified versions of domain-specific models, and there is room for improvements in both the inputs and the models used. An acknowledgement of the necessary balance needed between ease of use and detail in these models espouses further the model evaluation work described in the previous section. Integrated assessment models only require the least complex model that is suitable to the problem at hand, meaning model development and selection for future integrated assessments will rely on detailed evaluations like those in this work.

A potential future study is a comparison of future NO_X emissions controls and their impacts on ozone concentration distributions. This thesis showed that the low end of the range of ozone concentrations has increased just as high concentrations have decreased, but questions remain, for instance: *how will future emissions changes impact seasonal ozone distributions?* and *how will background ozone change in the future, and how will background levels impact future changes in ozone distributions?* These questions could be addressed using models discussed above and would leverage the United States' vast monitoring network and emissions inventory to develop spatial detrending and empirical emissions-concentration models. A study in this vein could take multiple approaches to comparing long-term, spatially broad variability in ozone with emissions and meteorology. For example, spatial applications of the detrending and statistical modelling techniques used in this thesis would allow for in-depth investigations into emissions and meteorological impacts on concentrations.

In chemical transport models, various tools provide estimates of how much an emitted pollutant contributes to simulated ambient concentrations (e.g., brute force methods, direct sensitivities estimates, and the Integrated Source Apportionment Method—ISAM) (Kwok et al. 2015). One thing that is achievable (but lacking), however, is a direct estimate of ozone production efficiency. This could be achieved by adding a 'counter' that sums the number of ozone molecules contributed by each NO_X molecule. Tested at various emission levels, this would give an idea for the OPE at future NO_X levels, and the optimal efficiency level achievable. Resolving ozone responses to emission changes at low concentrations is becoming important with the lowering national ambient air quality standard in the United States.

In summary, policy projections rely on experiences from accountability research and improvements in air quality models. Future work should prioritize model development and evaluation for specific applications, and should build on interdisciplinary collaborations with stakeholders in government, regulated industries, public health, and earth sciences.

9.3 References

Clancy, Luke, Pat Goodman, Hamish Sinclair, and Douglas W Dockery. 2002. "Effect of Air-Pollution Control on Death Rates in Dublin, Ireland: An Intervention Study." *Lancet* 360 (9341): 1210–14. doi:10.1016/S0140-6736(02)11281-5.

Cropper, Maureen, Arthur Fraas, and Richard Morgenstern. 2017. "Looking Backward to Move Forward." *Science* 355 (6332): 1375–76. http://www.sciencemagazinedigital.org/sciencemagazine/31_march_2017?sub_id=CDdH64IJVjiS&u1=01469606&folio=1375&pg=31#pg31.

de Hoogh, Kees, Michal Korek, Danielle Vienneau, Menno Keuken, Jaakko Kukkonen, Mark J. Nieuwenhuijsen, Chiara Badaloni, et al. 2014. "Comparing Land Use Regression and Dispersion Modelling to Assess Residential Exposure to Ambient Air Pollution for Epidemiological Studies." *Environment International* 73. Elsevier Ltd: 382–92. doi:10.1016/j.envint.2014.08.011.

Fiore, Arlene M., Vaishali Naik, Eric M. Leibensperger, and John D. Bachmann. 2015. "Air Quality and Climate Connections." *Journal of the Air & Waste Management Association* 65 (6): 645–85. doi:10.1080/10962247.2015.1040526.

Friberg, Mariel D., Xinxin Zhai, Heather A. Holmes, Howard H. Chang, Matthew J. Strickland, Stefanie Ebelt Sarnat, Paige E. Tolbert, Armistead G. Russell, and James A. Mulholland. 2016. "Method for Fusing Observational Data and Chemical Transport Model Simulations to Estimate Spatiotemporally Resolved Ambient Air Pollution." *Environmental Science and Technology* 50 (7): 3695–3705. doi:10.1021/acs.est.5b05134.

Friedman, MS, KE Powell, L Hutwagner, L M Graham, and W G Teague. 2001. "Impact of Changes in Transportation and Commuting Behaviors during the 1996 Summer Olympic Games in Atlanta on Air Quality and Childhood Asthma." *Jama* 285 (7): 897–905. http://www.ncbi.nlm.nih.gov/pubmed/11180733.

Health Effects Institute. 2003. "Assessing Health Impact of Air Quality Regulations: Concepts and Methods for Accountability Research." http://pubs.healtheffects.org/getfile.php?u=261.

Henneman, Lucas RF, Howard H Chang, David Lavoue, James A Mulholland, and Armistead G Russell. 2017. "Accountability Assessment of Regulatory Impacts on Ozone and PM2.5 Concentrations Using Statistical and Deterministic Pollutant Sensitivities." *Air Quality Atmosphere and Health*. doi:10.1007/s11869-017-0463-2.

Henneman, Lucas RF, Heather A Holmes, James A Mulholland, and Armistead G Russell. 2015. "Meteorological Detrending of Primary and Secondary Pollutant Concentrations: Method Application and Evaluation Using Long-Term (2000-2012) Data in Atlanta." *Atmospheric Environment* 119. Elsevier Ltd: 201–10. doi:10.1016/j.atmosenv.2015.08.007.

Henneman, Lucas RF, Cong Liu, Yongtao Hu, James A Mulholland, and Armistead G Russell. n.d. "Air Quality Modeling for Accountability Research: Operational, Dynamic, and Diagnostic Evaluation."

Henneman, Lucas RF, Cong Liu, James A Mulholland, and Armistead G Russell. 2016. "Evaluating the Effectiveness of Air Quality Regulations: A Review of Accountability Studies and Frameworks." *Journal of the Air & Waste Management Association* 67 (2). Taylor & Francis: 144–72. doi:10.1080/10962247.2016.1242518.

Henneman, Lucas RF, Peter Rafaj, Harold J. Annegarn, and Carmen Klausbruckner. 2015. "Assesing Emission Levels and Costs Associated with Climate and Air Pollution Policies in South Africa." *Energy Policy* 89. Elsevier: 160–70. doi:10.1016/j.enpol.2015.11.026.

Hu, Y., S. Balachandran, J. E. Pachon, J. Baek, C. Ivey, H. Holmes, M. T. Odman, J. A. Mulholland, and A. G. Russell. 2014. "Fine Particulate Matter Source Apportionment Using a Hybrid Chemical Transport and Receptor Model Approach." *Atmospheric Chemistry and Physics* 14 (11): 5415–31. doi:10.5194/acp-14-5415-2014.

Kloog, Itai, Francesco Nordio, Brent a Coull, and Joel Schwartz. 2012. "Incorporating Local Land Use Regression And Satellite Aerosol Optical Depth In A Hybrid Model Of Spatiotemporal PM2.5 Exposures In The Mid-Atlantic States." *Environmental Science & Technology*. doi:10.1002/dev.21214.Developmental.

Kwok, R. H F, K. R. Baker, S. L. Napelenok, and G. S. Tonnesen. 2015. "Photochemical Grid Model Implementation and Application of VOC, NOx, and O3 Source Apportionment." *Geoscientific Model Development* 8 (1): 99–114. doi:10.5194/gmd-8-99-2015.

National Academies of Sciences Engineering and Medicine. 2016. *The Future of Atmospheric Chemistry Research: Remembering Yesterday, Understanding Today, Anticipating Tomorrow*. Washington, DC: The National Academies Press. doi:10.17226/23573.

National Research Council. 2004. *Air Quality Management in the United States*. National Academies Press. doi:10.17226/10728.

Oreskes, Naomi, Kristin Shrader-Frechette, and Kenneth Belitz. 1994. "Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences." *Science* 263 (5147): 641–46.

Rich, David Q., Kaibo Liu, Jinliang Zhang, Sally W. Thurston, Timothy P. Stevens, Ying Pan, Cathleen Kane, et al. 2015. "Differences in Birth Weight Associated with the 2008 Beijing Olympic Air Pollution Reduction: Results from a Natural Experiment." *Environmental Health Perspectives* 117 (11): 1713–17. doi:10.1289/ehp.1408795.

Russell, Armistead G, Paige E Tolbert, Lucas RF Henneman, Joe Abrams, Cong Liu, Mitch Klein, James A Mulholland, et al. 2017. "Impacts of Regulations on Air Quality and Emergency Department visits in the Atlanta metropolitan area, 1999-2013" *In press*.

U.S. EPA. 2011. "The Benefits and Costs of the Clean Air Act from 1990 to 2020 Final Report."

APPENDIX A. SUPPLEMENTAL INFORMATION FOR CHAPTER 6. AIR QUALITY MODELING FOR ACCOUNTABILITY RESEARCH: OPERATIONAL, DYNAMIC, AND DIAGNOSTIC EVALUATION

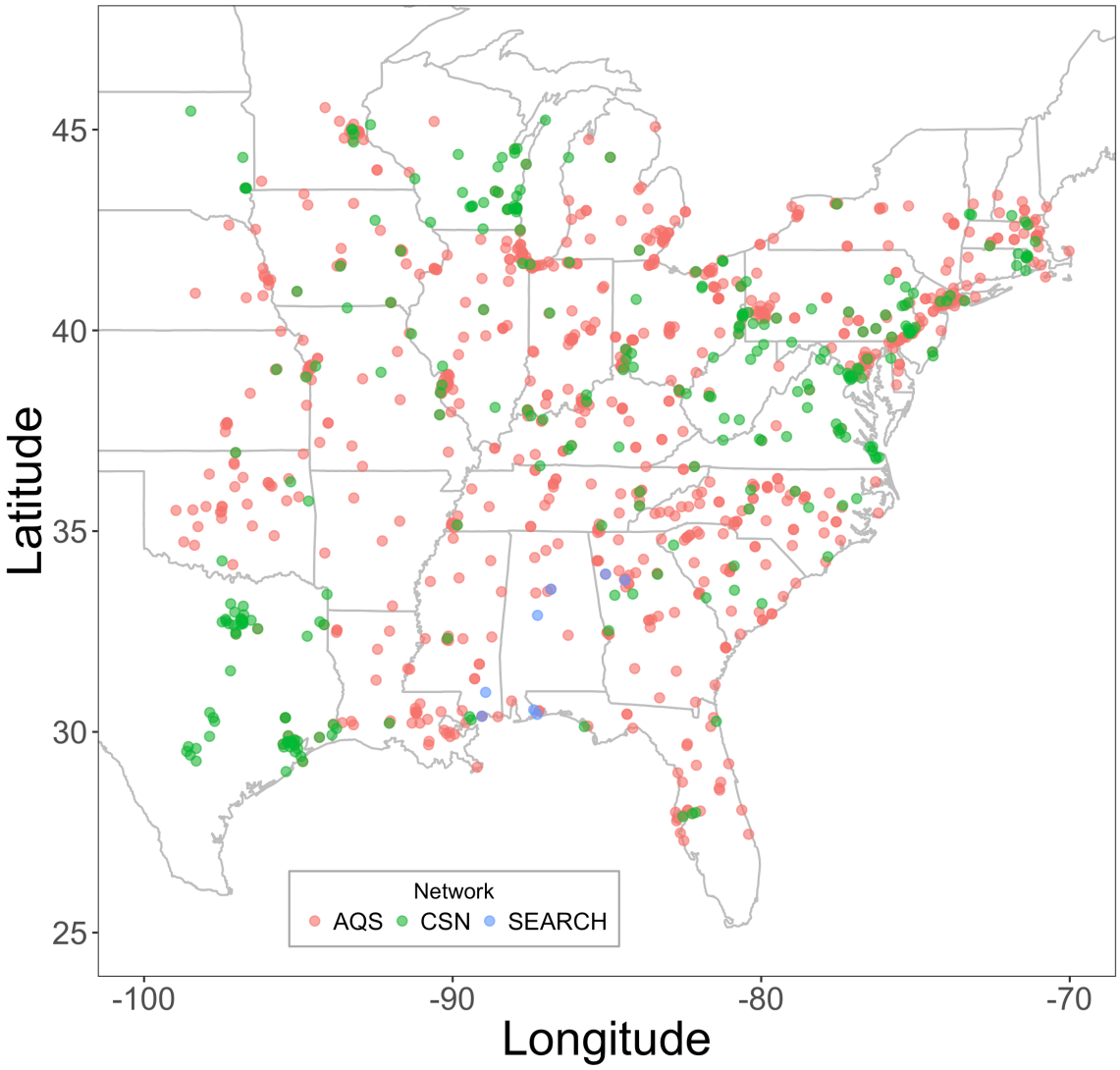


Figure A-1. Monitoring sites in the domain from three networks: EPA's Air Quality System (AQS) monitors, Chemical Speciation Network (CSN) and the SouthEastern Aerosol Research and Characterization (SEARCH) network. 8 SEARCH sites are located in the Southeast as four urban-rural pairs.

Table A-1. Detailed CMAQ model evaluation for all SEARCH monitoring sites for, cumulative, 2001, 2002, 2011 and 2012. Hourly ozone is evaluated with a cutoff of 40 ppb except for r.

Performance Metric/Year	Ozone (hourly)	Ozone MDA8h	$PM_{2.5}$	Ozone MDA8h	$PM_{2.5}$
N	222487	9664	9176	3875	3449
NMB	5.0%*	-7.2%*	21%†	-3.5%†	7.0%*
NME	8.0%*	16%*	46%*	74%*	86%*
r	0.63*	0.71*	0.48*	0.68	0.53*

Table A-2. CMAQ model valuation for four years (2001, 2002, 2011, and 2012) at JST. EGU and MOB columns compare empirical sensitivities to CMAQ-modeled sensitivities. The SO_4^{2-} and EC empirical models do not estimate MOB and EGU sensitivities, respectively.

	Concentration			EGU			MOB		
	NMB (%)	NME (%)	r	NMB (%)	NME (%)	r	NMB (%)	NME (%)	r
Gaseous									
O_3 (MDA8h)	4.0†	19.4*	0.82†	-26.3	82.4	0.59*	-30.9	73.4	0.59*
NO_2	324.2‡	324.2‡	0.40‡	-40.2‡	110.2‡	-0.05‡	89.5‡	107.5‡	0.04‡
CO	154.7‡	155.2‡	0.62‡	-89.3‡	98.1‡	0.05‡	62.6‡	73.6‡	0.54‡
SO_2	-60‡	66.2‡	0.56‡	-81.4‡	81.9‡	0.24‡	-69.1‡	71.6‡	0.37‡
Particulate									
$PM_{2.5}$	-4.7†	35.2*	0.52*	-46.2	57.5	0.57*	-5.3†	53.1	0.21
SO_4^{2-}	-18.1*	46.5*	0.49*	-16.2*	58.4	0.27	—	—	—
NH_4^+	-15.1*	52.1	0.40	-38.7	70.5	0.48*	-48.1	68.0	0.13
NO_3^-	40.3*	80.6*	0.68‡	-123.4	133.6	-0.19‡	529.6	822.1	-0.05‡
EC	51.1	65.5*	0.60‡	—	—	—	-35.4*	39.0†	0.63‡
OC	-34.2*	43.8*	0.49‡	228.8	308.3	-0.01‡	-53.3	98.7	0.02‡

* Meets the criteria benchmark
† Meets the goal benchmark
‡ No benchmark available

A.1 Operational Evaluation

EC and OC

Figure A-2 illustrates a consistent trend of overestimation for EC throughout 2001 and 2011. Performance is improved in 2011 compared to 2001, probably as a result of better characterization of biomass burning and, possibly, mobile emissions. While R is close to or higher than 0.4 (no criteria value for R is proposed for PM species by Emery et al.) in 2011, NMB and NME exceeded their criteria bar (+/- 40%, 75%) in January (88%, 103%), February (114%, 120%), March (103%, 115%), April (66%, 82%), November (63%, 86%), and December (68%, 96%) when biomass burning leads to large emissions of EC. It should be noted that EC and OC are operationally defined. Thermal optical reflectance (TOR) was used for 2011 in AQS and thermal optical transmittance (TOT) was used for 2001. Values of EC and OC for 2001 were converted to TOR method for a more consistent evaluation between the years (Malm et al. 2011).

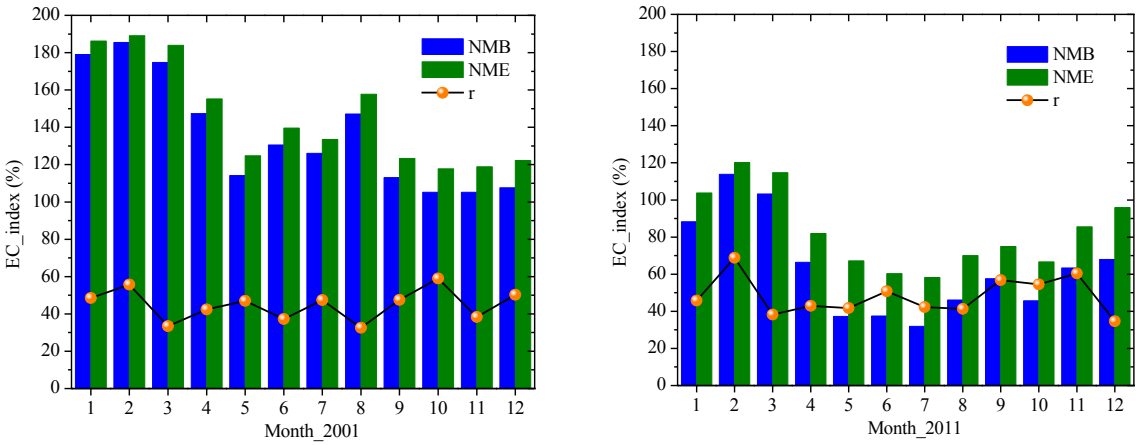


Figure A-2. Operational results for EC. Left: 2001, right: 2011. R is presented here as a percent; i.e., the result from Eq. 6.3 has been multiplied by 100.

CMAQ performance on OC exhibits a similar monthly pattern to $PM_{2.5}$ mass, overestimating in cold months and underestimating in warm months. Comparing 2011 to 2001, improvement in R is observed, especially for the three months with lowest R in 2001, February, April, and November. Correlation coefficients (r) are increased from 0.14, 0.035, and 0.098 to 0.56, 0.16, and 0.58 for these three months between 2001 and 2011, respectively. NMB and NME, however, increase in January, February, and December, and exceed criteria values (+/- 50, 65%, respectively). The overestimation in cold months might

be attributed to biomass burning, as OC and EC usually share the same sources. Secondary organic aerosol (SOA), a substantial component of OC, has been well documented to be underestimated in summer (Carlton et al. 2010; Pye et al. 2013).

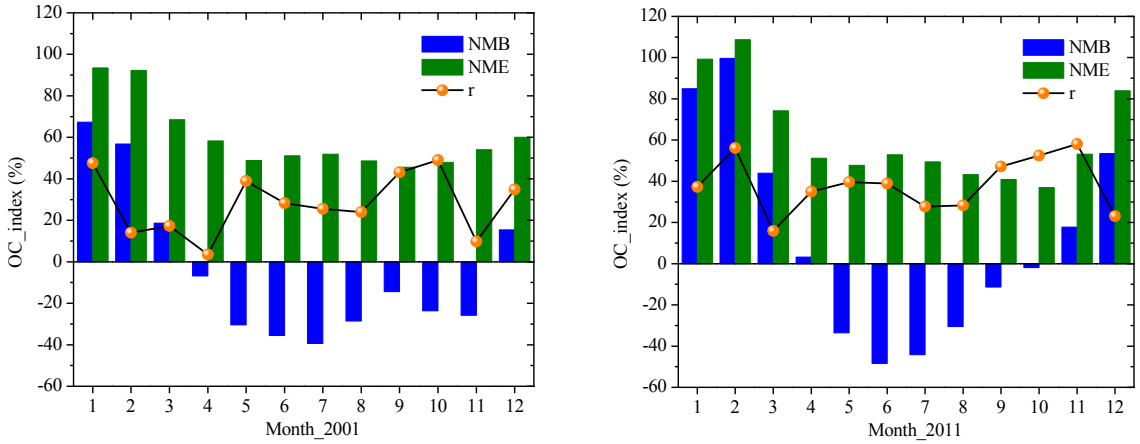


Figure A-3. Operational evaluation results for OC. Left: 2001, right: 2011. R is presented here as a percent; i.e., the result from Eq. 6.3 has been multiplied by 100.

Sulfate, nitrate, and ammonium

CMAQ tends to underestimate sulfate from January to September, and overestimate from October and December, as shown in Figure A-4. Koo et al. (2015) reported an underestimating trend for sulfate throughout the year. NMB seems significantly reduced between 2011 and 2001, especially in highest biased months of May (-38% in 2001 and -24% in 2011), June (-46% in 2001 and -23% in 2011), and August (-51% in 2001 and -40% in 2011). NMB in July remains similar in both years (-40% in 2001 and -44% in 2011). Sulfate generally meets the criteria for NMB and NME (+/- 30% and 50%, respectively) in 2011, except in July and August. Recent studies suggest that mobile source, NOx emission estimates may be biased high by a factor of about two (Anderson et al. 2014; Goldberg et al. 2016; Travis et al. 2016). This might, at least partially, explain the underestimation, as NOx competes with SO2 for oxidants, though uncertainty in precipitation and associated wet scavenging has also been cited as a potential reason (Xing et al. 2015).

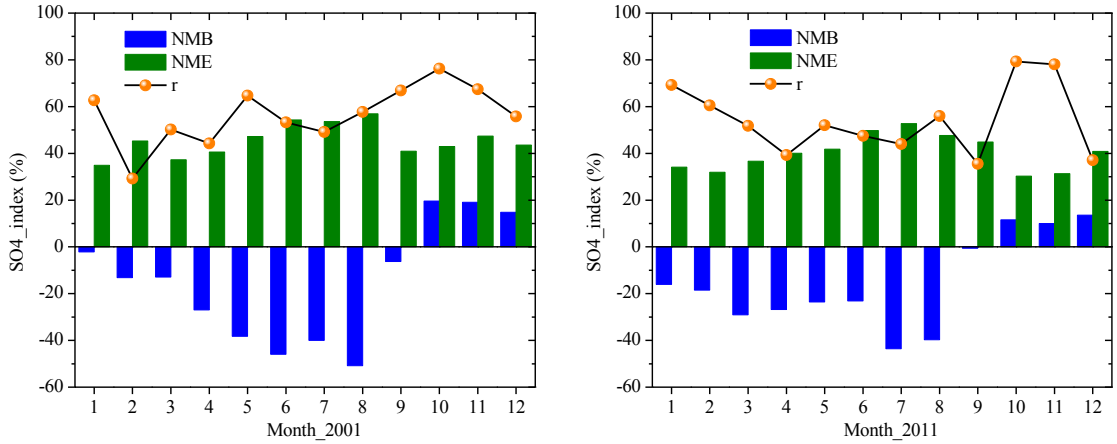


Figure A-4. Operational evaluation results for sulfate. Left: 2001, right: 2011. R is presented here as a percent; i.e., the result from Eq. 6.3 has been multiplied by 100.

Figure A-5 shows that NMB and NME for nitrate in 2001 and 2011 are consistently smaller than +/- 65%, and 115%, their criteria values respectively. CMAQ overestimates nitrate in cold months like February to April, October, and November, and underestimates in warm months such as July and August. The high bias in NOx emissions, if that is the case, would lead to overestimation high bias in nitrate. Nitrate concentrations are low in summer time, and the underestimate in July and August is small on an absolute basis and has a minor influence on annual averaged modeled $PM_{2.5}$ mass concentrations. NME decreases from 2001 to 2011, especially in June, October, and November.

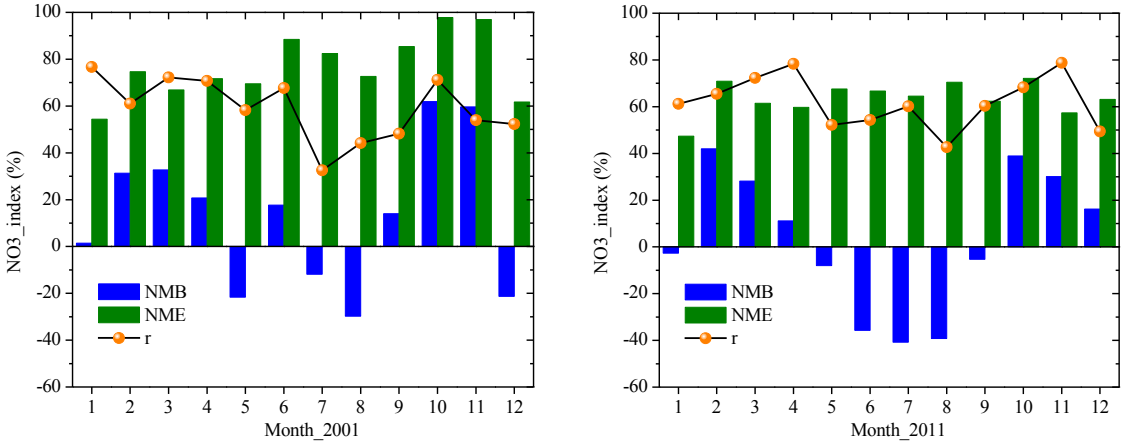


Figure A-5. Operational evaluation results for nitrate. Left: 2001, right: 2011. R is presented here as a percent; i.e., the result from Eq. 6.3 has been multiplied by 100.

Modeled ammonium has a low bias through January to May, as indicated in Figure A-6. NMB is reduced in May, June, August, October, and November, from 2001 (-27%, -

31%, -43%, 74%, 61%) to 2011 (-5.4%, -18%, -32%, 46%, 50%). Despite this, criteria of NMB and NME(+/- 30 and 50%, respectively) are still violated in August (-32% and 50%), October (46% and 60%), and November (50% and 64%). Ammonium concentration is comparatively low in cold months, as emission of its main sources, fertilizer and animal feedlots, depends on temperature. Underestimation in June, July, and August might be related to underestimation of sulfate, as these two species track each other closely (Weber et al. 2016). Correlation coefficient (r) is generally higher than 0.4, although no criteria is suggested for PM species. This means that CMAQ can track ammonium variability well.

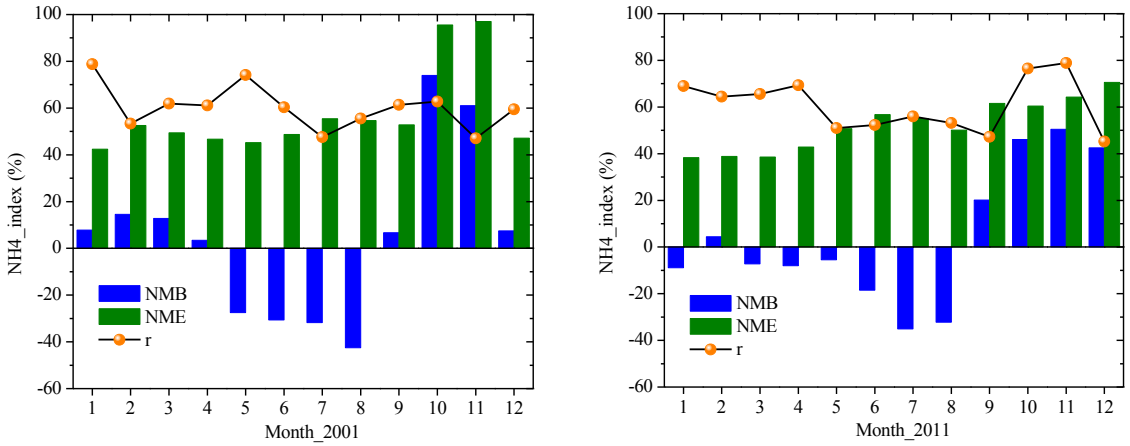


Figure A-6. Operational performance results for ammonium. Left: 2001, right: 2011. R is presented here as a percent; i.e., the result from Eq. 6.3 has been multiplied by 100.

A.2 Dynamic evaluation - JST

We focus our dynamic evaluation in the supplemental on the comparison between 2001 and 2011 at JST.

$PM_{2.5}$ mass

We calculated NMB, NME, and *R* to evaluate modeled change of $PM_{2.5}$ between 2001 and 2011 in comparison to observed change at JST. As mentioned before, due to data scarcity, dynamic evaluation was conducted on a seasonal basis. Modeled PM concentration changes yield a NMB_{Δ} of -13% to 6.6%, indicating a good performance of CMAQ to capture seasonally averaged change of $PM_{2.5}$ from 2001 to 2011. In contrast, NME falls in the range of 53% (in fall) to 86% (in summer). This means CMAQ simulates $PM_{2.5}$ concentration changes with a high error, probably due to NME of $PM_{2.5}$ concentration consistently higher than 40% in 2001 and 2011. R is higher than 0.6 in winter, spring, and fall, and is 0.35 in summer. These results imply that work is needed

directed at improving CMAQ's ability to better capture $PM_{2.5}$ concentration changes in summer.

$PM_{2.5}$ species

Normalized comparison values (NMB_{Δ} and NME_{Δ}) for $PM_{2.5}$ species tend to be larger than what is found for the operational evaluation. This phenomenon which is driven in part by using the differences between values in 2011 and 2001, which are smaller than the actual observations (Table A-3). For example, in the extreme case, if the values did not change between 2001 and 2011, the normalized metrics would be undefined (infinity). The change in sulfate concentration is overestimated throughout the seasons, especially in winter and spring. Change in nitrate is underestimated in winter, spring, summer, but is overestimated in fall. Changes of ammonium and OC are negatively biased in all four seasons. EC is overestimated in winter, spring, and summer, but is underestimated in fall. Overall, results suggest CMAQ simulates temporal variability of change of PM species relatively well. When interpreting the results, recall that nitrate concentrations in summer and fall tend to be low, and sulfate levels in winter and spring tend to be low.

Table A-3. Dynamic performance of CMAQ to capture $PM_{2.5}$ species changes between 2001 and 2011 for using observations only at the Jefferson Street, Atlanta, Georgia, US SEARCH site.

Metric	Season	Sulfate	Nitrate	Ammonium	EC	OC
NMB_{Δ}	Winter	146%	-143%	-60.0%	62.7%	-72.0%
	Spring	55.1%	-29.8%	-40.3%	27.6%	-46.5%
	Summer	6.0%	-56.1%	-63.4%	6.8%	-0.3%
	Fall	10.6%	79.6%	-41.8%	-34.1%	-80.1%
NME_{Δ}	Winter	156%	158%	86.6%	102%	79.8%
	Spring	110%	92.9%	63.9%	142%	78.0%
	Summer	82.3%	66.0%	71.4%	129%	89.6%
	Fall	74.6%	180%	56.0%	67.3%	103%
r	Winter	0.64	0.51	0.22	0.68	0.83
	Spring	0.43	0.42	0.39	0.28	0.67
	Summer	0.40	0.28	0.48	0.34	0.25
	Fall	0.61	0.64	0.56	0.57	0.59

Sulfate (+), OC (-), and nitrate (-) dominate mean bias of $PM_{2.5}$ concentration change between 2001 and 2011 in winter. The signs in the parentheses indicate directions of biases. In spring, change of sulfate remains biased high, and contributes more than the other species, while OC (-) and ammonium (-) are non-negligible. Ammonium (-) drives

modeling bias in summer. MB of change of OC is lower than 0.1 ug/m3 in summer, although it is significantly underestimated in both years (Figure A-7). In fall, MB of change of OC (-) is up to -2.7 µg m^{-3}, probably partially explained by opposite bias in September (underestimation in 2001 and overestimation in 2011).

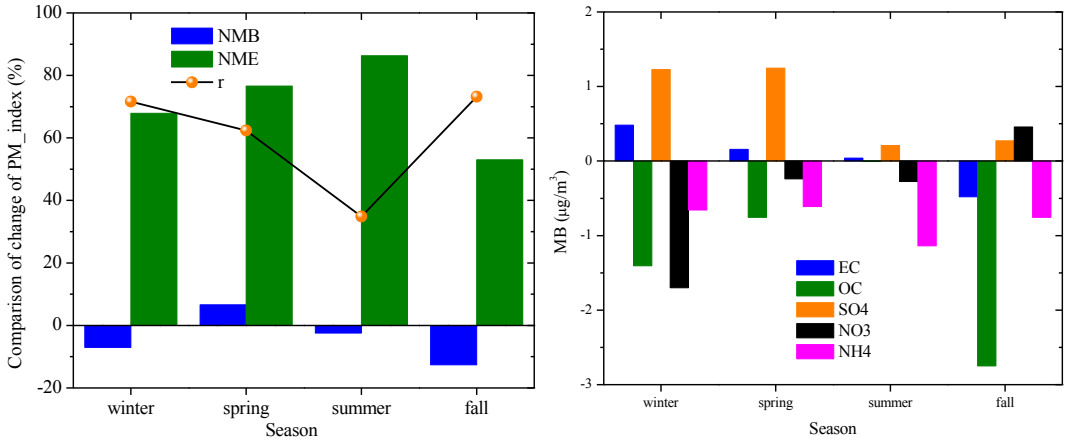


Figure A-7. Results for the dynamic evaluation of $PM_{2.5}$ between 2001 and 2011, using observations only at the Jefferson Street, Atlanta, Georgia, US SEARCH site (left). Mean Bias (MB) for difference of $PM_{2.5}$ species between 2001 and 2011 at Jefferson Street, Atlanta, Georgia, US (right). R is presented here as a percent; i.e., the result from Eq. 6-8 has been multiplied by 100.

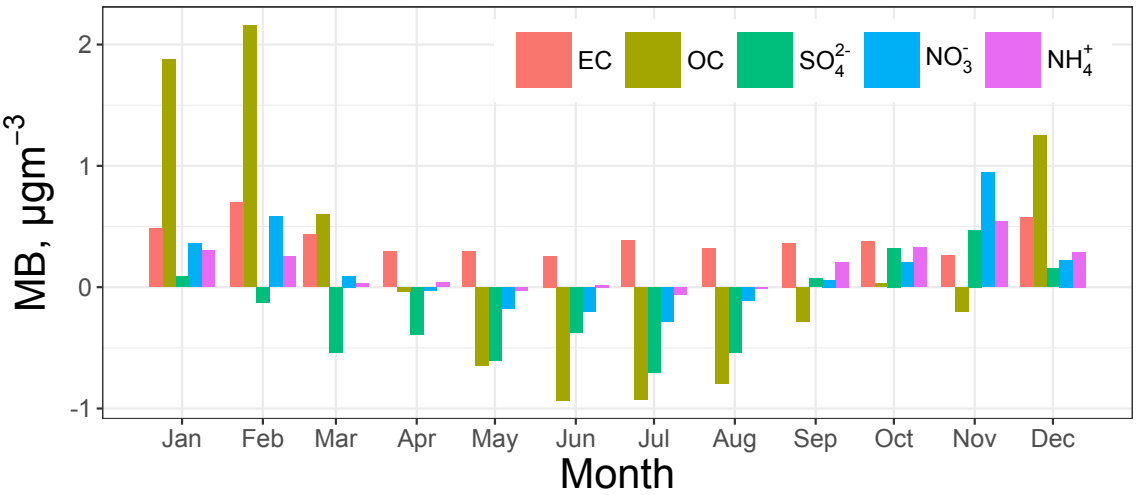


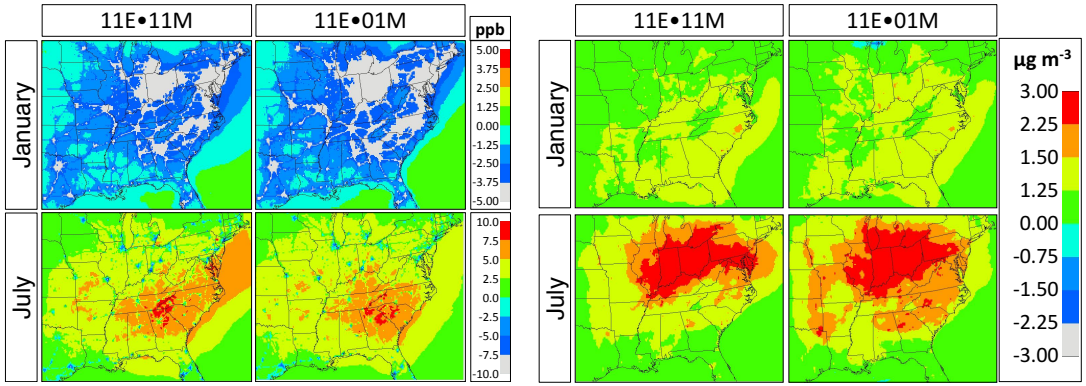
Figure A-8. Results for the dynamic evaluation of $PM_{2.5}$ species between 2002 and 2012, using observations across all available CSN monitors.

Table A-4. Operational evaluation for all modeling years in two sub-domains: 'North' and 'South' of 37° latitude. Results are reported for all available AQS sites.

		2001		2002		2011		2012	
		North	**South**	**North**	**South**	**North**	**South**	**North**	**South**
O_3	N	75048	62526	67019	78164	107417	82581	62854	47358
	NMB	2.7	8.0	9.6	1.1	5.8	2.7	-0.24	4.3
	NME	20	21	23	19	18	18	15	18
	r	0.70	0.67	0.7	0.75	0.67	0.67	0.72	0.65
$PM_{2.5}$	N	23192	9320	23623	12304	22724	13954	22494	13173
	NMB	16	-4.5	8.1	-18.5	5.2	-21.2	11.9	-15.2
	NME	49	46	50	46	49	50	49	48
	r	0.55	0.30	0.46	0.29	0.40	0.19	0.50	0.27

Table A-5. Domain-wide CMAQ Model Performance for ozone in 2011 for two cases: the base case and with mobile NO_x emissions reduced by half. Hourly ozone is evaluated with a cutoff of 40 ppb except for r. All AQS sites are used.

Performance Metric/Year	Ozone Hourly	Ozone MDA8h
11M:11E (July)		
N	485557	20037
NMB	-10.09	14.5%
NME	22.7%	22.0%
R	0.66	0.53
11M:11E (July, reduced mobile NO_X)		
N	485557	20037
NMB	-13.0%	9.3%
NME	22.5%	20.3%
R	0.66	0.50



Ozone *$PM_{2.5}$*

Figure A-9. Monthly-averaged sensitivities for ozone (left) and $PM_{2.5}$ to on-road emissions in two dynamic runs.

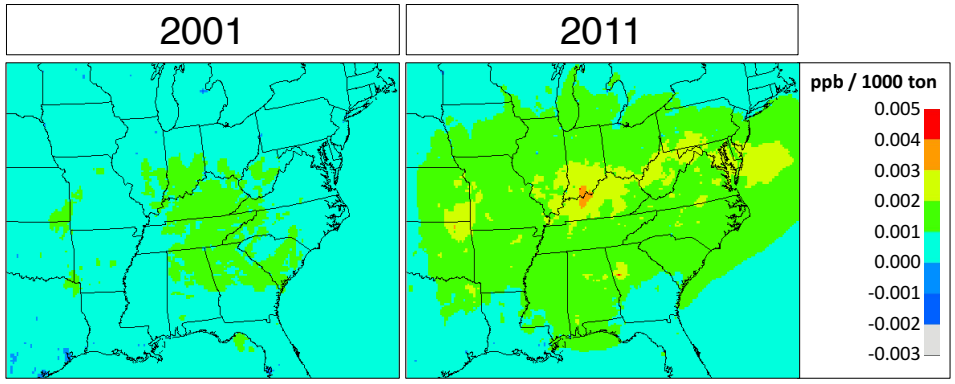


Figure A-10. EGU sensitivities normalized by total EGU NO_X emissions in the domain in each year. Note that there is no special heterogeneity in the normalization, but that, overall, per unit sensitivities have increased.

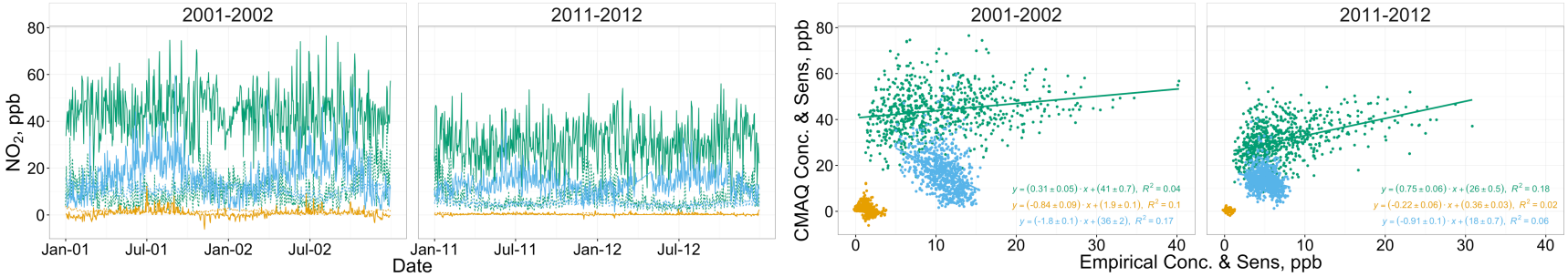


Figure A-11. NO_2 daily CMAQ-empirical comparisons (time series on the left, scatter plot on the right) for ambient concentrations, mobile, and EGU sensitivities.

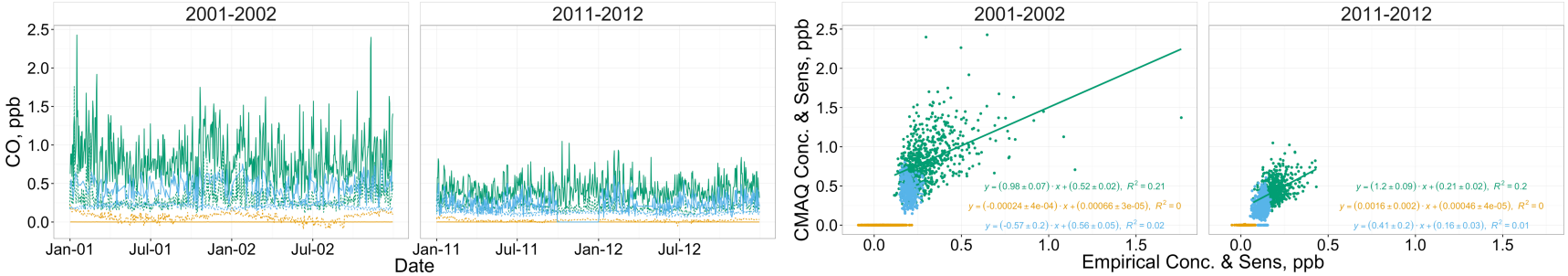


Figure A-12. CO daily CMAQ-empirical comparisons (time series on the left, scatter plot on the right) for ambient concentrations, mobile, and EGU sensitivities.

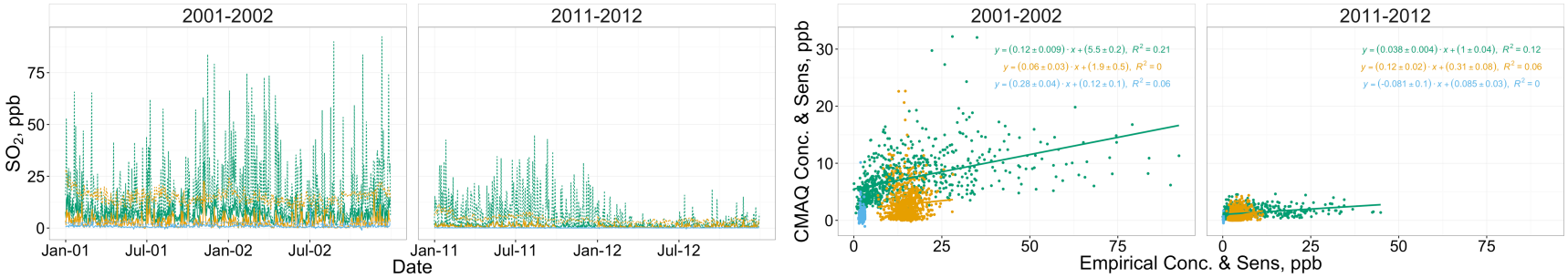


Figure A-13. SO_2 daily CMAQ-empirical comparisons (time series on the left, scatter plot on the right) for ambient concentrations, mobile, and EGU sensitivities.

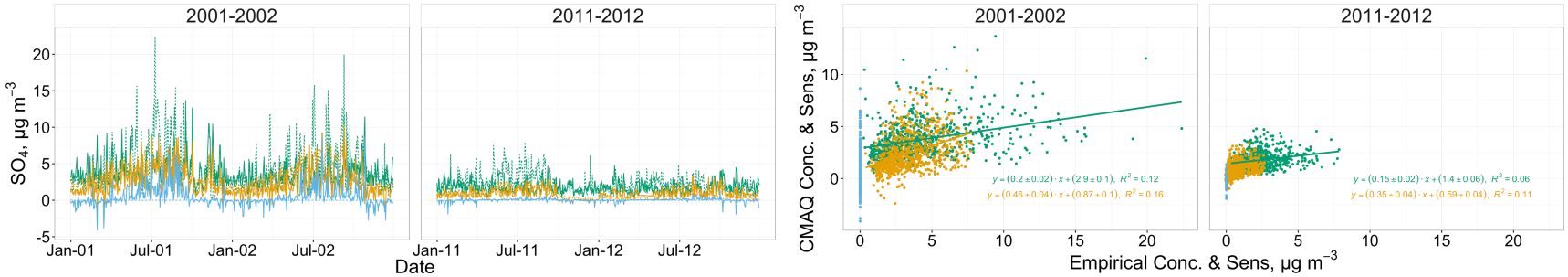


Figure A-14. SO_4^{2-} daily CMAQ-empirical comparisons (time series on the left, scatter plot on the right) for ambient concentrations, mobile, and EGU sensitivities.

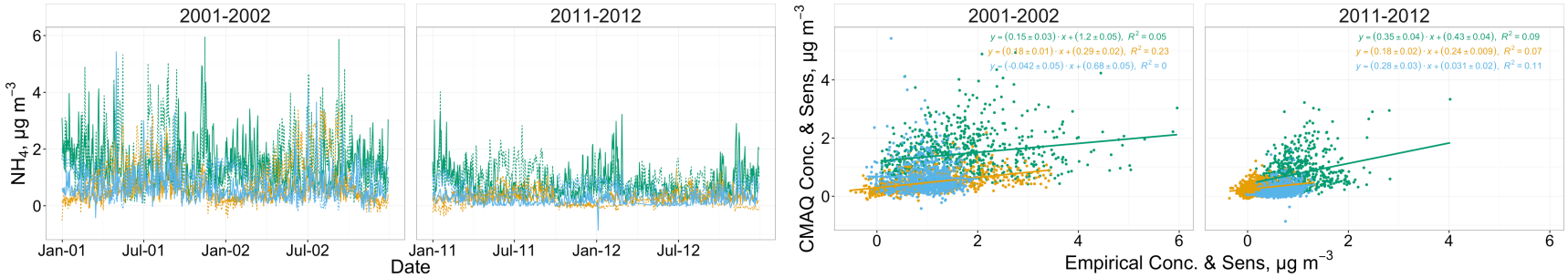


Figure A-15. NH_4^+ daily CMAQ-empirical comparisons (time series on the left, scatter plot on the right) for ambient concentrations, mobile, and EGU sensitivities.

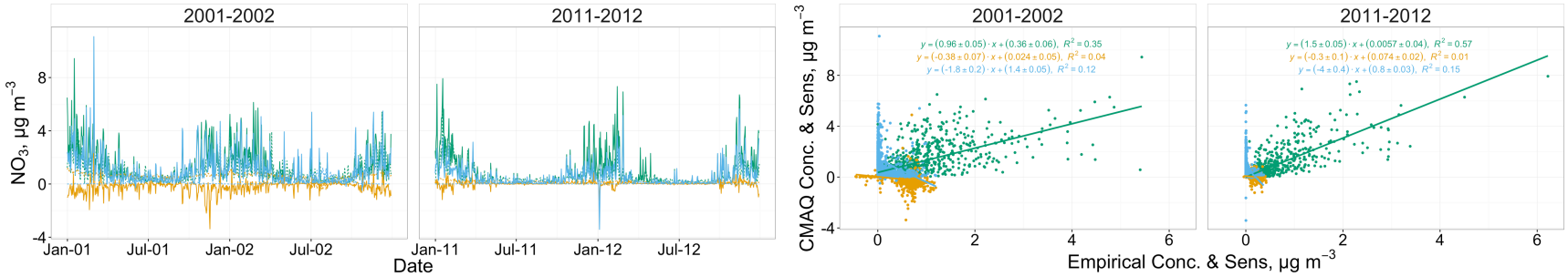


Figure A-16. NO_3^- daily CMAQ-empirical comparisons (time series on the left, scatter plot on the right) for ambient concentrations, mobile, and EGU sensitivities.

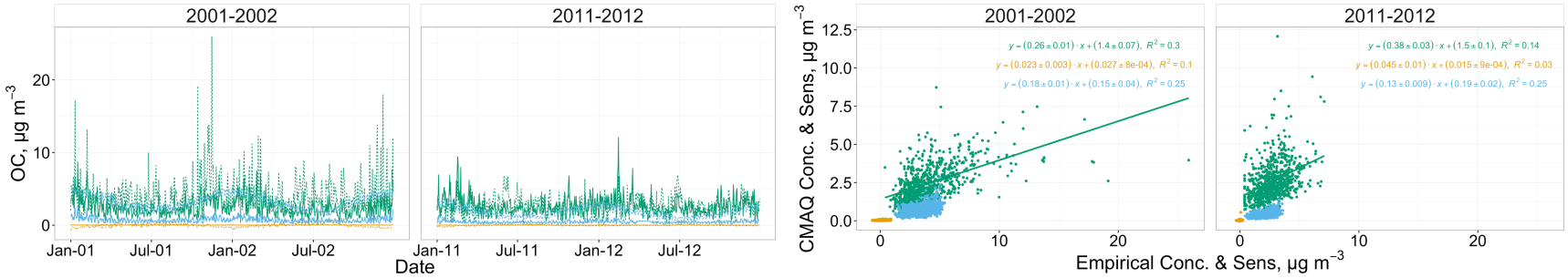


Figure A-17. OC daily CMAQ-empirical comparisons (time series on the left, scatter plot on the right) for ambient concentrations, mobile, and EGU sensitivities.

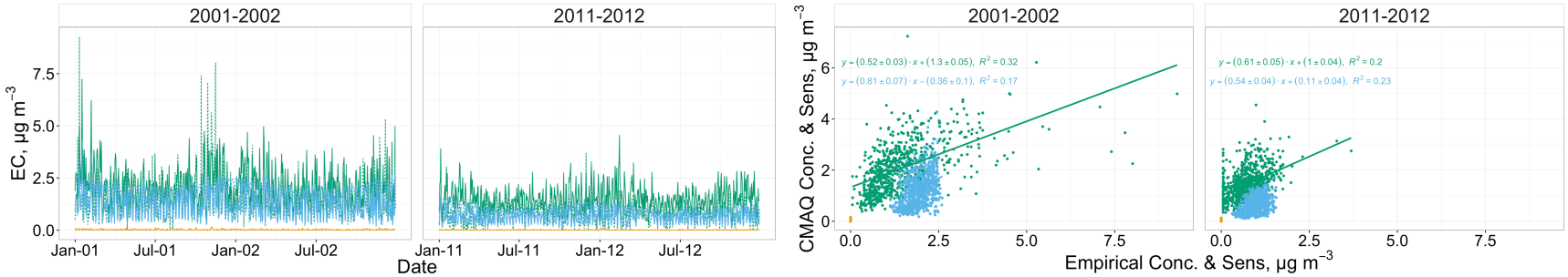


Figure A-18. EC daily CMAQ-empirical comparisons (time series on the left, scatter plot on the right) for ambient concentrations, mobile, and EGU sensitivities.

Table A-6. Average spatial and temporal correlations. No cutoffs were used.

Performance metric/year	O_3 (hourly)	O_3 (MDA8h)	$PM_{2.5}$
2001			
Spatial		0.56	0.44
Temporal		0.69	0.47
2002			
Spatial	0.46	0.57	0.42
Temporal	0.68	0.74	0.39
2011			
Spatial		0.52	0.39
Temporal		0.67	0.42
2012			
Spatial	0.39	0.53	0.41
Temporal	0.63	0.69	0.47

A.3 References

Anderson DC, Loughner CP, Diskin G, Weinheimer A, Canty TP, Salawitch RJ, et al. 2014. Measured and modeled CO and NOy in DISCOVER-AQ: An evaluation of emissions and chemistry over the eastern US. Atmos. Environ. 96:78–87; doi:10.1016/j.atmosenv.2014.07.004.

Carlton AG, Bhave P V, Napelenok SL, Edney EO, Sarwar G, Pinder RW, et al. 2010. Model representation of secondary organic aerosol in CMAQv4.7. Environ. Sci. Technol. 44:8553–60; doi:10.1021/es100636q.

Goldberg DL, Vinciguerra TP, Anderson DC, Hembeck L, Canty TP, Ehrman SH, et al. 2016. CAMx Ozone Source Attribution in the Eastern United States using Guidance from Observations during DISCOVER-AQ Maryland. Geophys. Res. Lett. 2249–2258; doi:10.1002/2015GL067332.

Koo B, Kumar N, Knipping E, Nopmongcol U, Sakulyanontvittaya T, Odman MT, et al. 2015. Chemical transport model consistency in simulating regulatory outcomes and the relationship to model performance. Atmos. Environ. 116:159–171; doi:10.1016/j.atmosenv.2015.06.036.

Malm WC, Schichtel BA, Pitchford ML. 2011. Uncertainties in PM2.5 Gravimetric and Speciation Measurements and What We Can Learn from Them. J. Air Waste Manage. Assoc. 61:1131–1149; doi:10.1080/10473289.2011.603998.

Pye HOT, Pinder RW, Piletic IR, Xie Y, Capps SL, Lin YH, et al. 2013. Epoxide pathways improve model predictions of isoprene markers and reveal key role of acidity in aerosol formation. Environ. Sci. Technol. 47:11056–11064; doi:10.1021/es402106h.

Travis KR, Jacob DJ, Fisher JA, Kim PS, Marais EA, Zhu L, et al. 2016. Why do models overestimate surface ozone in the Southeast United States? Atmos. Chem. Phys. 16:13561–13577; doi:10.5194/acp-16-13561-2016.

Weber RJ, Guo H, Russell AG, Nenes A. 2016. High aerosol acidity despite declining atmospheric sulfate concentrations over the past 15 years. Nat. Geosci. 9:1–5; doi:10.1038/NGEO2665.

Xing J, Mathur R, Pleim J, Hogrefe C, Gan CM, Wong DC, et al. 2015. Observations and modeling of air quality trends over 1990-2010 across the Northern Hemisphere: China, the United States and Europe. Atmos. Chem. Phys. 15:2723–2747; doi:10.5194/acp-15-2723-2015.

APPENDIX B. SUPPLEMENTAL INFORMATION FOR CHAPTER 7. OZONE RESPONSES TO RECENT AND FUTURE EMISSIONS CHANGES IN THE EASTERN UNITED STATES

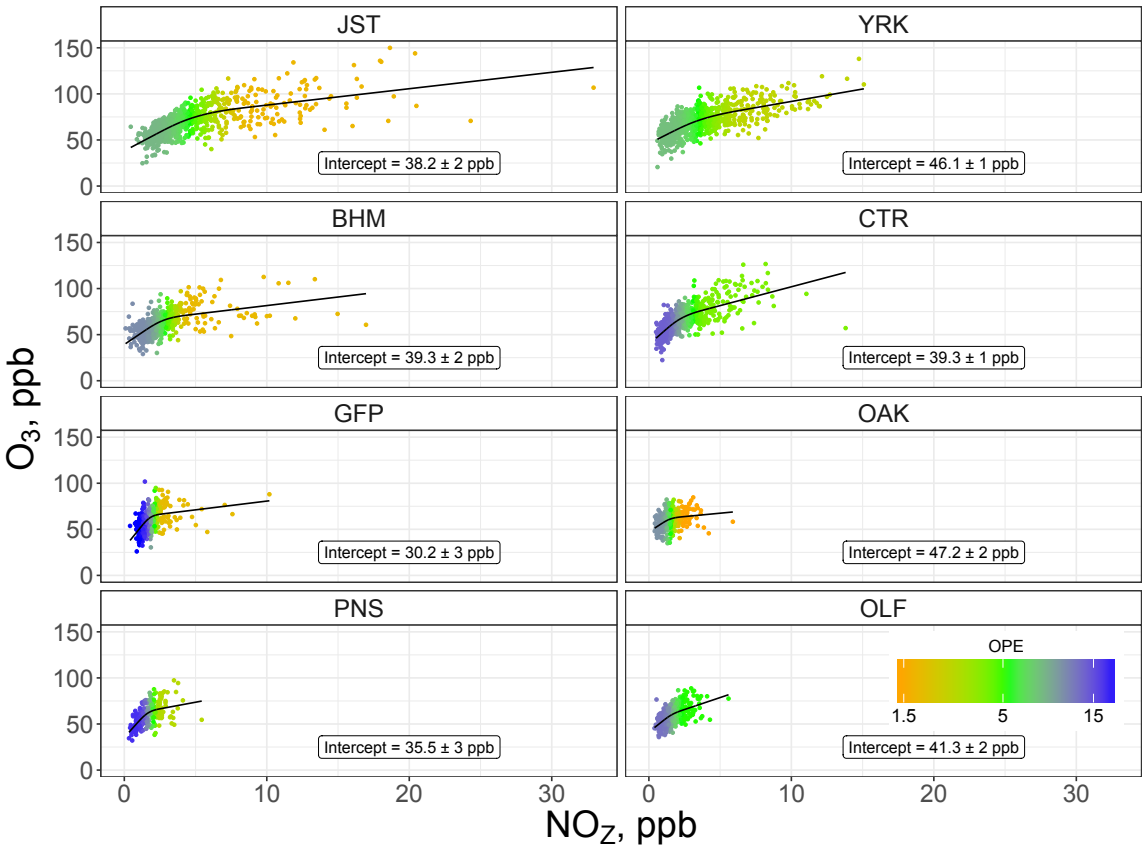


Figure B-1. Empirical OPE fits using natural spline with 3 knots. Urban sites are on the left, and rural/suburban sites are on the right; each urban/rural (or suburban) pair is matched right to left. Points are colored by the slope of the calculated curve at that NO_Z level (i.e., their OPE). Models are fit for days with highest 20% PS.*

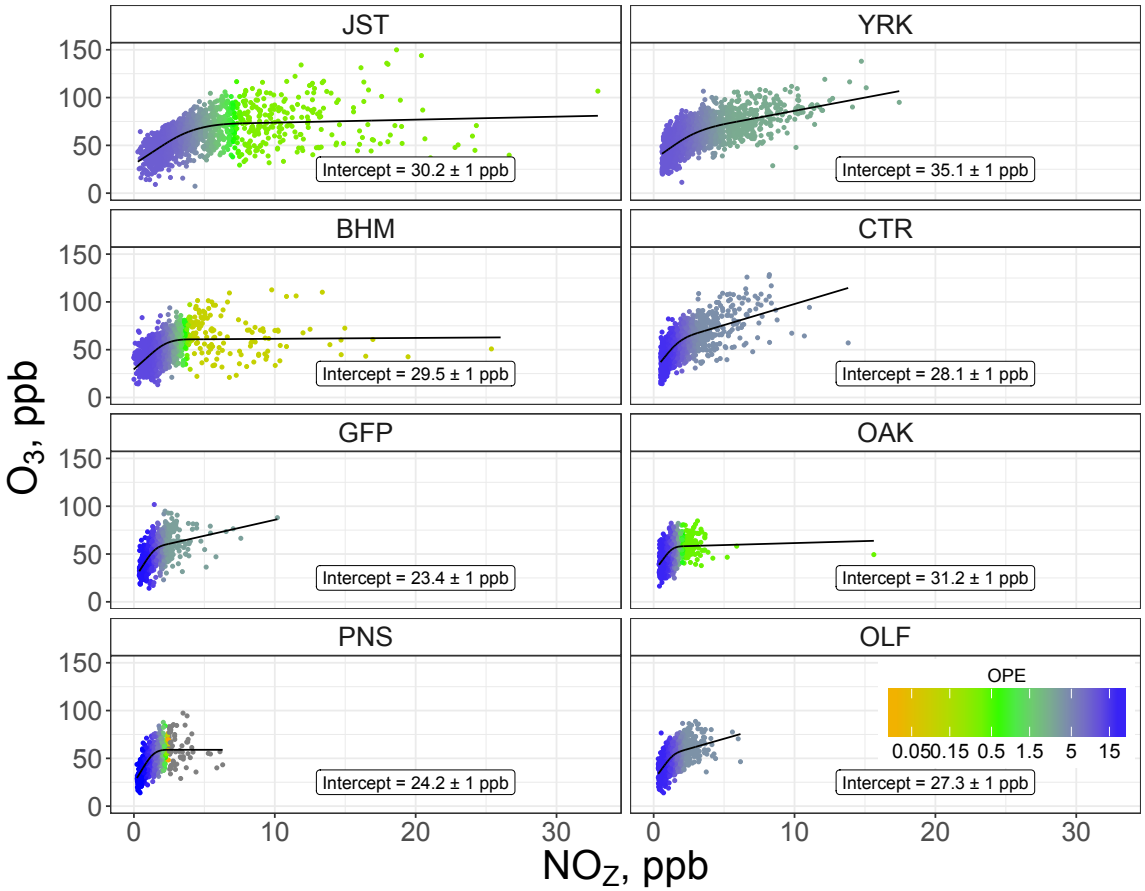


Figure B-2. OPE fit for 50% of days (instead of 20%).

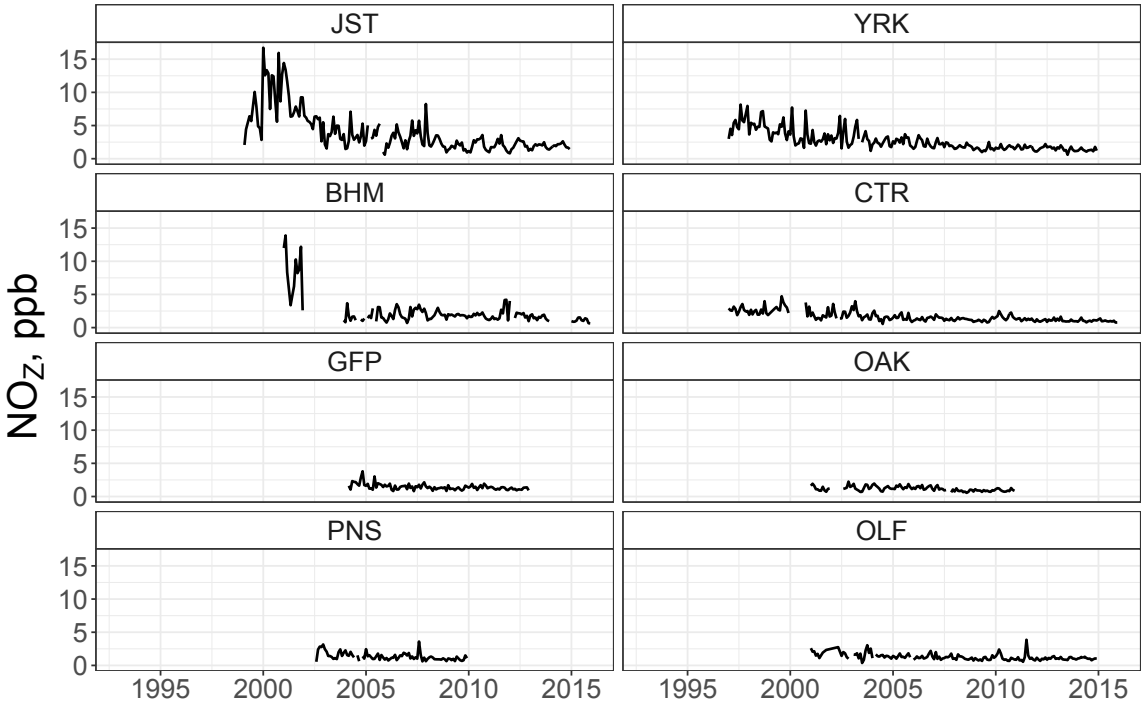


Figure B-3. Monthly-averaged NO_Z concentrations at each SEARCH site.

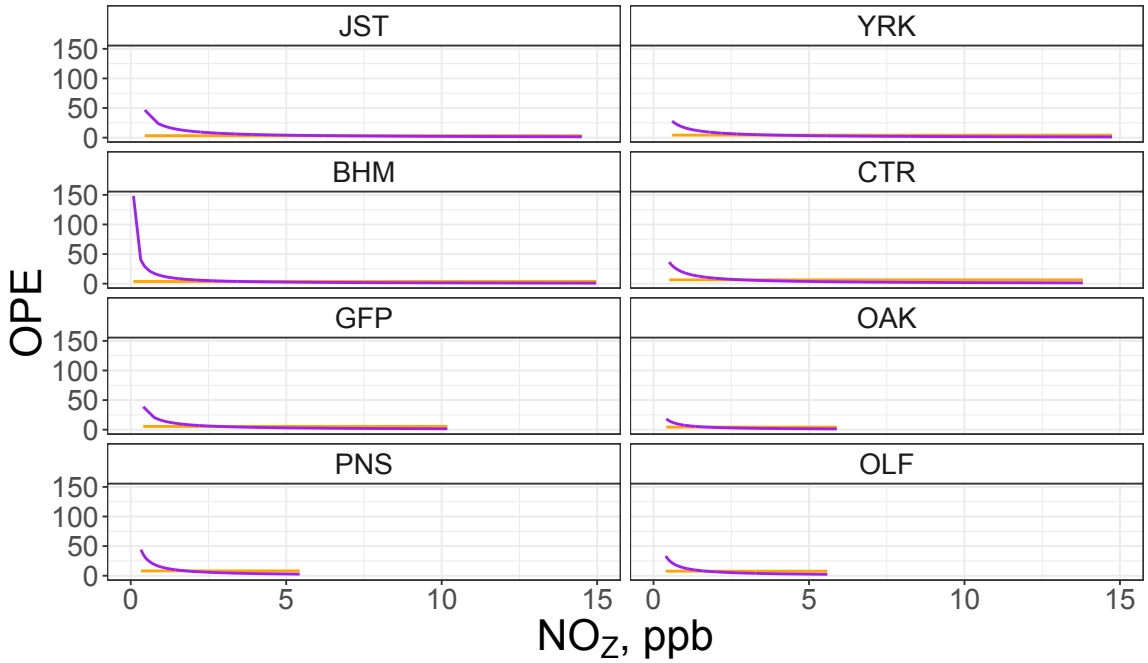


Figure B-4. OPEs estimated using linear (orange) and log (purple) models fit to the O_3 vs. NO_Z relationship (Figure B-1).

Table B-1. July OPEs estimated via the deposition-corrected sensitivity method from CMAQ-DDM output.

		Mean OPE (std. error)			
Site	**Category**	**2001**	**2011**	**2011 50% NOx**	**2011 10% NOx**
JST	Urban	3.9	7.8	9.5	11.4
YRK	Rural	10.6	12.0	13.0	13.8
BHM	Urban	3.2	6.8	9.3	11.2
CTR	Rural	14.4	15.9	16.0	15.4
GFP	Urban	8.7	18.9	17.8	17.3
OAK	Rural	13.6	16.1	16.6	15.5
PNS	Urban	3.2	12.8	13.8	14.8
OLF	Suburban	1.0	7.6	10.4	12.4

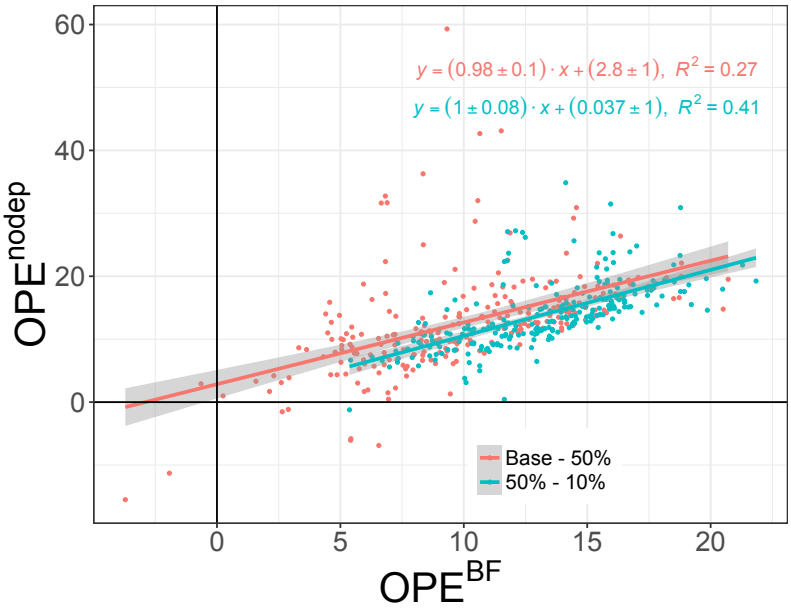


Figure B-5. Comparison of OPE brute force with sensitivity (not including deposition). Error values reported in the equation are the standard error.

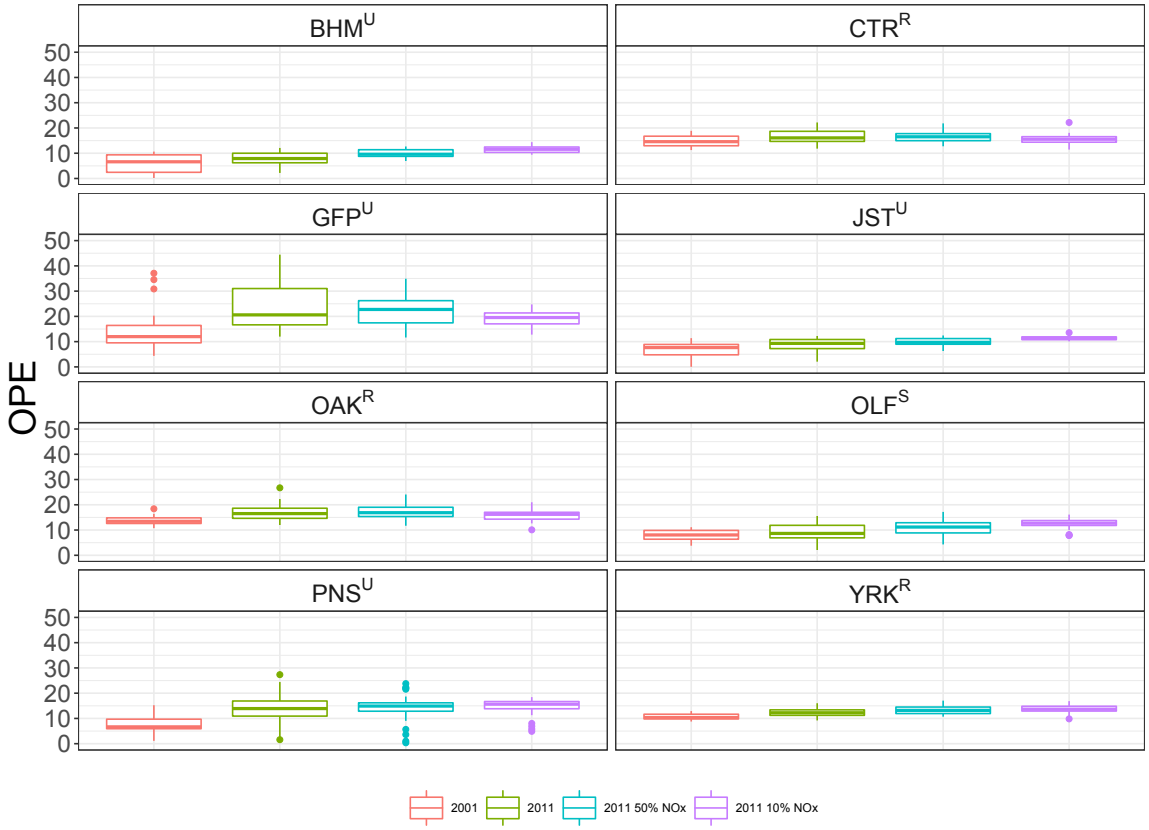


Figure B-6. July OPEs estimated via the deposition-corrected sensitivity method from CMAQ-DDM output.

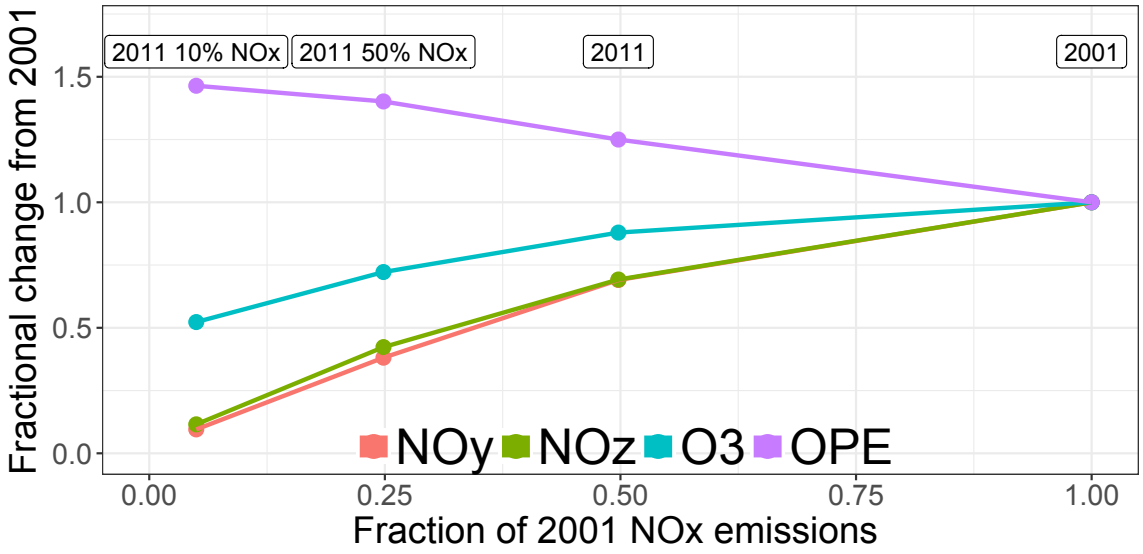


Figure B-7. Fractional changes in NO_Z, NO_Y, O_3, and OPE with reducing emissions. Overland CMAQ simulations (information in absolute terms in Table 7-1).

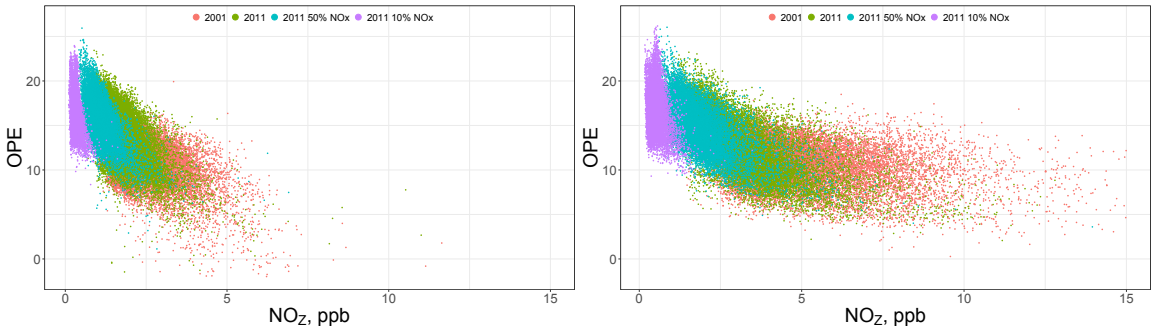


Figure B-8. Left graph: All over-land locations for average July. Right: All over-land locations for days selected with 2nd highest O_3 by location.

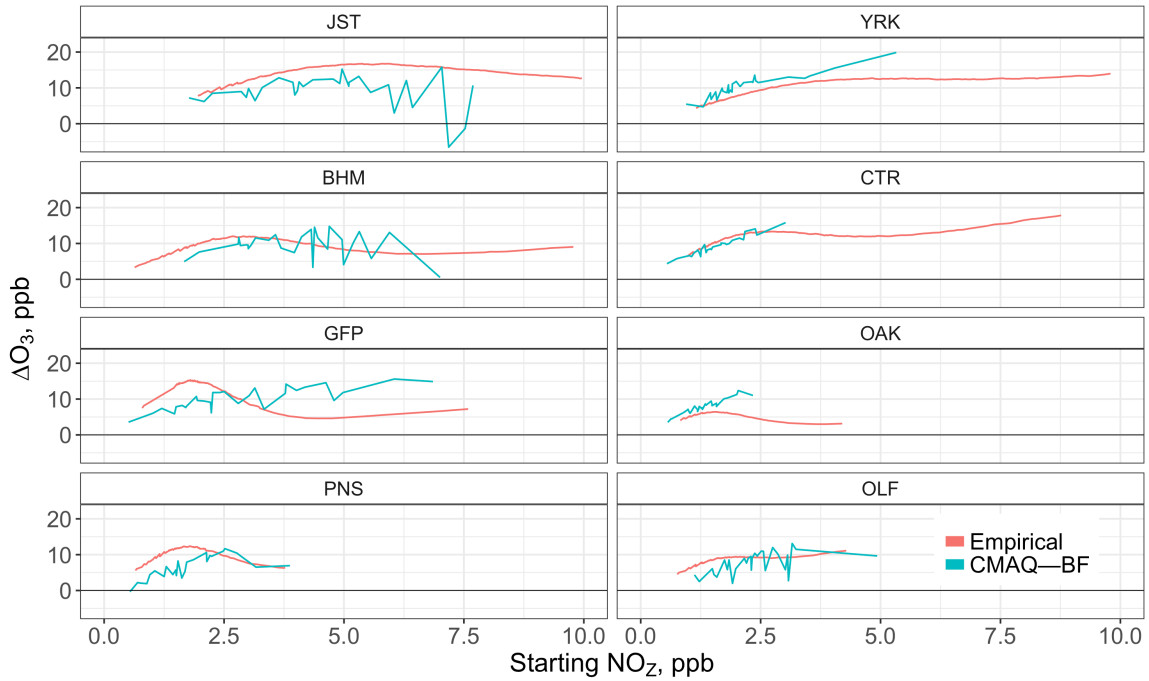


Figure B-9. Change in CMAQ-simulated O_3 for 50% reduction in NO_Z – some sites suggestive of multiple effectiveness regimes, similar to finding using empirical method.

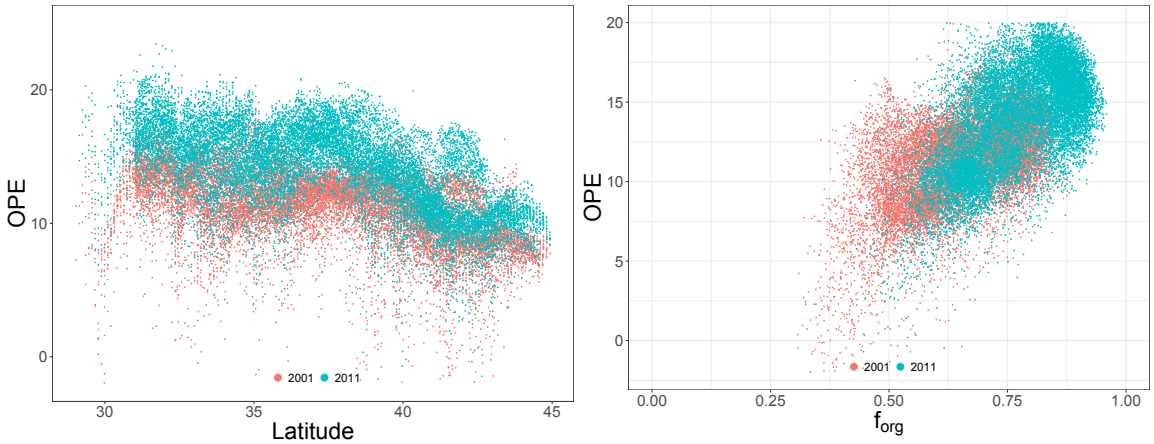


Figure B-10. CMAQ domain-wide comparisons between latitude and the fraction of nitrogen in organic molecules (f_{org}).

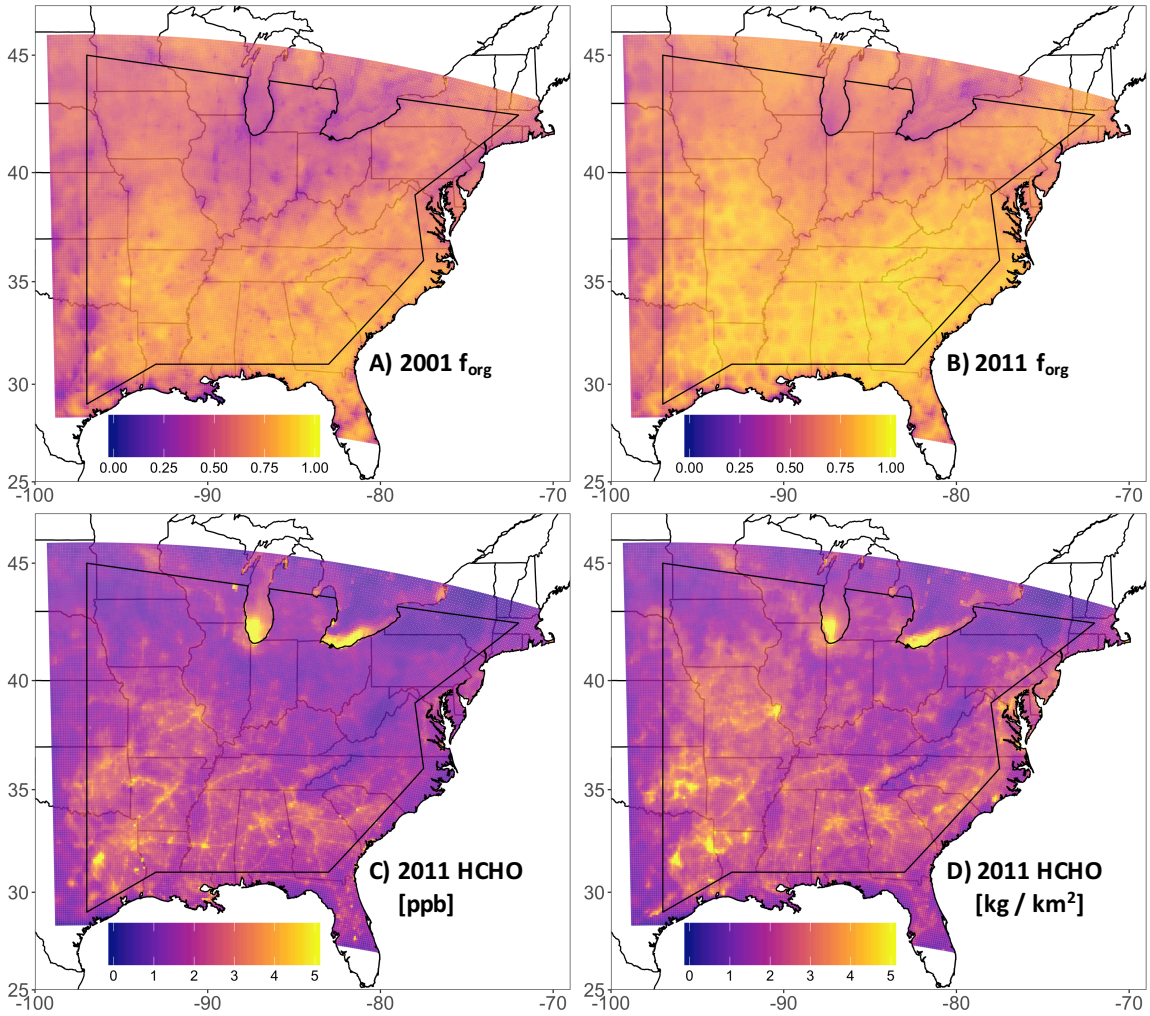


Figure B-11. A&B: plots of fraction of NO_Z that is organic (PAN+NTR+PANX)/NO_Z. C&D: formaldehyde. All are July average 2-3 p.m.

Table B-2. Regression summaries for average July OPE and latitude (Lat), NO_X, NO_Z, HCHO, and the fraction of NO_Z in organic molecules ($f_{N,org}$).

Year	R^2	Intercept (Lat = 29)	Lat	NO_X	NO_Z	HCHO	$f_{N,org}$
2001	0.16	14.0 (0.05)	-0.34 (0.01)	---	---	---	---
2001	0.63	14.9 (0.04)	-0.30 (0.00)	-0.80 (0.01)	---	---	---
2001	0.39	18.1 (0.07)	-0.28 (0.01)	---	-1.75 (0.02)	---	---
2001	0.29	20.1 (0.12)	-0.56 (0.01)	---	---	-2.19 (0.04)	---
2001	0.26	3.9 (0.22)	-0.09 (0.01)	---	---	---	12.26 (0.26)
2001	0.64	14.7 (0.23)	-0.32 (0.01)	-0.73 (0.01)	-0.02 (0.03)	-0.64 (0.03)	2.32 (0.24)
2011	0.30	18.0 (0.05)	-0.48 (0.01)	---	---	---	---
2011	0.56	19.1 (0.04)	-0.44 (0.00)	-1.27 (0.01)	---	---	---
2011	0.52	21.6 (0.06)	-0.34 (0.00)	---	-2.54 (0.03)	---	---
2011	0.39	22.5 (0.11)	-0.59 (0.01)	---	---	-1.58 (0.03)	---
2011	0.39	5.5 (0.26)	-0.23 (0.01)	---	---	---	13.39 (0.27)
2011	0.60	16.2 (0.26)	-0.35 (0.01)	-0.85 (0.02)	-0.74 (0.04)	-0.49 (0.03)	5.34 (0.27)

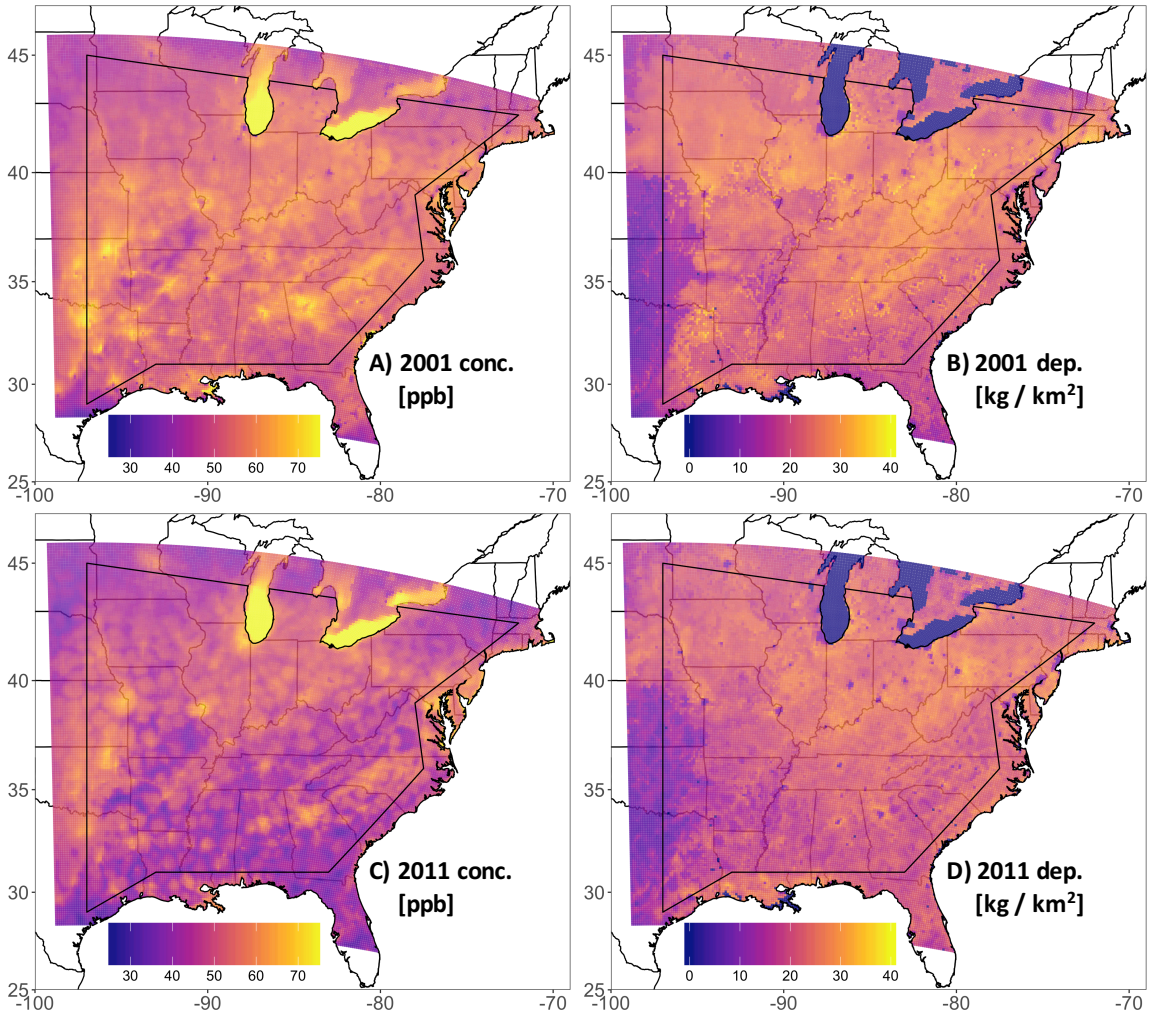


Figure B-12. Ozone concentration and deposition (July 2-3 p.m. average).

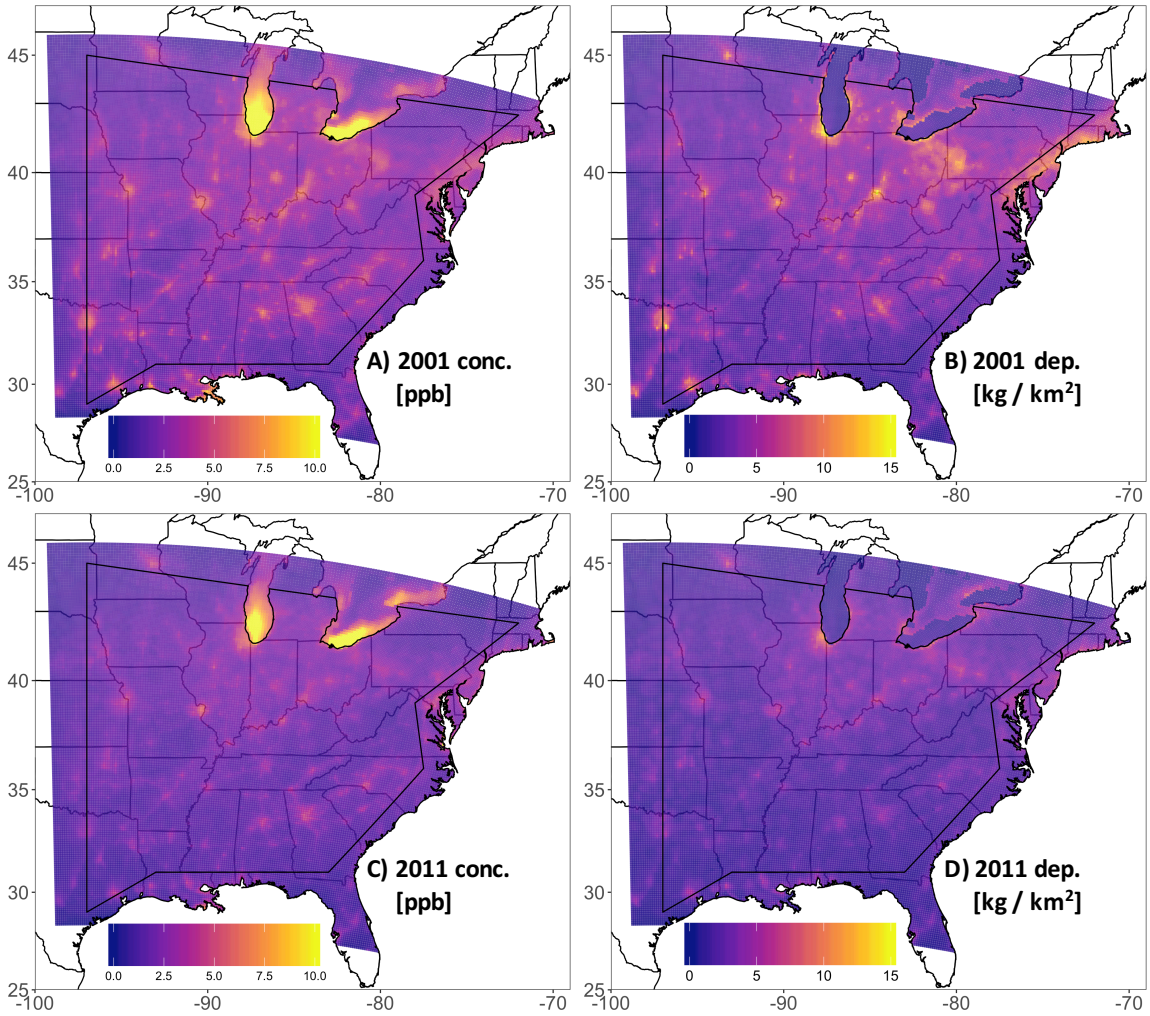


Figure B-13. NO_Z concentration and deposition (July 2-3 p.m. average).